

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 05:04:17 ; Search time 24 Seconds

(without alignments)  
296.681 Million cell updates/sec

Title: US-09-684-725-2

Perfect score: 1263

Sequence: 1 MEKLNASWYQOKLEDPQ.....LLPMTVISVLYMALRVSI 242

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

#### SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	1252	99.1	415	4	US-09-545-944-2
2	304	24.1	353	1	US-08-118-270-45
3	304	24.1	353	1	PCT-US93-08528-45
4	281	22.2	416	3	US-08-858-876A-4
5	281	22.2	416	4	US-09-472-880-4
6	277	21.9	289	4	US-09-077-675A-10
7	277	21.9	361	4	US-09-077-675A-8
8	277	21.9	366	4	US-09-077-675A-13
9	274	21.7	353	4	US-09-077-675A-3
10	274	21.7	364	4	US-09-077-675A-16
11	273	21.6	289	4	US-09-077-675A-5
12	265	21.0	393	1	US-07-629-104I-3
13	262.5	20.8	259	4	US-09-261-599B-3
14	262.5	20.8	259	4	US-09-456-455A-3
15	259	20.5	398	2	US-08-288-663A-1
16	258	20.4	410	3	US-08-858-876A-2
17	258	20.4	410	4	US-09-472-880-2
18	256.5	20.3	400	4	US-09-351-198-2
19	256.5	20.3	400	4	US-09-113-426-2
20	256.5	20.3	415	4	US-08-405-271A-20
21	255.5	20.2	400	4	US-08-188-275A-2
22	253.5	20.1	400	4	US-08-889-108-8
23	253.5	20.1	400	5	PCT-US94-10358-8
24	252.5	20.0	369	1	US-07-816-283-6
25	252.5	20.0	369	1	US-08-417-103-6
26	252.5	20.0	369	1	US-08-417-103-16
27	247.5	19.6	369	2	US-08-411-859-3

28	247.5	19.6	369	4	US-08-387-707-9	Sequence 9, Appl 1
29	247.5	19.6	369	4	US-08-405-271A-9	Sequence 9, Appl 1
30	247.5	19.6	398	3	US-09-170-331-5	Sequence 5, Appl 1
31	246	19.5	410	4	US-09-200-090-2	Sequence 2, Appl 1
32	245.5	19.4	369	4	US-08-120-601B-9	Sequence 9, Appl 1
33	242.5	19.2	356	4	US-08-430-286A-2	Sequence 2, Appl 1
34	242.5	19.2	356	4	US-08-430-286A-5	Sequence 2, Appl 1
35	242.5	19.2	367	4	US-08-405-271A-23	Sequence 23, Appl 1
36	242.5	19.2	391	2	US-08-454-549-3	Sequence 3, Appl 1
37	242.5	19.2	391	2	US-08-454-552-3	Sequence 3, Appl 1
38	242.5	19.2	398	1	US-08-149-093A-5	Sequence 5, Appl 1
39	242.5	19.2	398	2	US-08-911-245-5	Sequence 5, Appl 1
40	242.5	19.2	398	3	US-08-889-108-2	Sequence 2, Appl 1
41	242.5	19.2	398	4	US-08-120-601B-2	Sequence 2, Appl 1
42	242.5	19.2	398	4	US-08-188-275A-3	Sequence 3, Appl 1
43	242.5	19.2	398	4	US-08-387-707-16	Sequence 16, Appl 1
44	242.5	19.2	398	4	US-09-510-473-5	Sequence 5, Appl 1
45	242.5	19.2	398	4	US-09-351-198-3	Sequence 3, Appl 1

#### ALIGNMENTS

```
RESULT 1
US-09-545-944-2
; Sequence 2, Application US/09545944
; Patent No. 6461836
; GENERAL INFORMATION:
; APPLICANT: AMES, ROBERT
; APPLICANT: ELSHOURBAGY, NABIL
; APPLICANT: MICHALOVICH, DAVID
; APPLICANT: SARAU, HENRY
; APPLICANT: SHARON, USMAN
; APPLICANT: WAMTER, LISA
; TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM RECEPTOR
; FILE REFERENCE: GP70657-1
; CURRENT APPLICATION NUMBER: US/09/545,944
; CURRENT FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: US 09/435,384
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 415
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-545-944-2
Query Match          99.1%; Score 1252; DB 4; Length 415;
Best Local Similarity 99.6%; Pred. No. 3e+107;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 MEKLNASWYQOKLEDPQKHNSTEYLAFPGRRSHFPLVSVYVPIEVGVIGN 60
    |||||||
Db 4 MEKLNASWYQOKLEDPQKHNSTEYLAFPGRRSHFPLVSVYVPIEVGVIGN 63
OY 61 VIVCVLVIHQAKTPNNYLFSLASDLVLLGMPLEYEMKRNPLFGVGCYFKT 120
    |||||||
Db 64 VIVCVLVIHQAKTPNNYLFSLASDLVLLGMPLEYEMKRNPLFGVGCYFKT 123
OY 121 ALPEYVCFASILSTTVSVRYVAIIHPFRAKLOSTRRLRLIGIYGFVSFLPNTS 180
    |||||||
Db 124 ALPEYVCFASILSTTVSVRYVAIIHPFRAKLOSTRRLRLIGIYGFVSFLPNTS 183
OY 181 IHGIRKHYFPNGSLVPSACCTVIRKPMIYNFTIOVTSFLFLLPMTVISVLYMALRV 240
    |||||||
Db 184 IHGIRKHYFPNGSLVPSACCTVIRKPMIYNFTIOVTSFLFLLPMTVISVLYMALRV 243
RESULT 2
US-08-118-270-45
; Sequence 45, Application US/08118270
```

```

; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; NUMBER OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-118-270-45

Query Match      24.1%; Score 304; DB 1; Length 353;
Best Local Similarity 33.3%; Pred. No. 1.9e-20;
Matches 68; Conservative 49; Mismatches 73; Indels 14; Gaps 5;

QY 45 VSVVVYPIFVGVYGNVLCVILQH---QAMKPTNYLFLSLVSDLLVLLGMFLVY 101
DB 4 VTAIYIALFVYGVGNSVTAFTLARKKSLQSLQSVHYHLSDLSLSDLLILM---VELY 60
DB 102 EMKRNYPFLGPGVC--YFKTALFETVCASLITVSVERYVALILHFRKLOSTR 158
DB 61 NFIWHHPMAFGDAGCGYF--LRDCTYATALNVAASLSVERVYLAICHPFKAKTLMRS 117
QY 159 RALRIIGIYWGVSFLSPNTSIHGKIFHYFPNGSLVPGSATCTYIKPMIYFNFIQVTS 218
DB 118 RTKKFISALIMLASALAIPLMFLTLGLQNR--SGDGTNPGSLVCTPIVDATYAKVVIQVNT 175
QY 219 FLEYLPMTVISVLYYIMALRVS I 242
DB 176 FMSFLPMLVISTLNTVINKLTV 199

RESULT 3
PCT-US93-08528-45
; Sequence 45, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; NUMBER OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK

```

```

; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-08528-45

Query Match      24.1%; Score 304; DB 5; Length 353;
Best Local Similarity 33.3%; Pred. No. 1.9e-20;
Matches 68; Conservative 49; Mismatches 73; Indels 14; Gaps 5;

QY 45 VSVVVYPIFVGVYGNVLCVILQH---QAMKPTNYLFLSLVSDLLVLLGMFLVY 101
DB 4 VTAIYIALFVYGVGNSVTAFTLARKKSLQSLQSVHYHLSDLSLSDLLILM---VELY 60
DB 102 EMKRNYPFLGPGVC--YFKTALFETVCASLITVSVERYVALILHFRKLOSTR 158
DB 61 NFIWHHPMAFGDAGCGYF--LRDCTYATALNVAASLSVERVYLAICHPFKAKTLMRS 117
QY 159 RALRIIGIYWGVSFLSPNTSIHGKIFHYFPNGSLVPGSATCTYIKPMIYFNFIQVTS 218
DB 118 RTKKFISALIMLASALAIPLMFLTLGLQNR--SGDGTNPGSLVCTPIVDATYAKVVIQVNT 175
QY 219 FLEYLPMTVISVLYYIMALRVS I 242
DB 176 FMSFLPMLVISTLNTVINKLTV 199

RESULT 4
US-08-858-876A-4
; Sequence 4, Application US/08858876A
; Patent No. 6028656
; GENERAL INFORMATION:
; APPLICANT: Daniel CAPUT
; APPLICANT: Pascual CHALON
; APPLICANT: Pascual FERRARA
; APPLICANT: VITA NATALIO
; TITLE OF INVENTION: TYPE 2 Neurotensin Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
; STREET: 400 Seventh Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004

```



RESULT 8  
 US-09/077-675A-13  
 Sequence 13 Application US/09077675A  
 Patent No. 6421199  
 GENERAL INFORMATION:  
 APPLICANT: Pai, Lee-Yuh  
 APPLICANT: Feigner, Scott C.  
 APPLICANT: Howard, Andrew D.  
 APPLICANT: Peng, Sheng-Shung  
 APPLICANT: Van Der Ploeg, Leonardus H. T.  
 TITLE OF INVENTION: RECEPTOR ASSAY  
 NUMBER OF SEQUENCES: 16  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Merck & Co., Inc.  
 STREET: P.O. Box 2000, 126 E. Lincoln Ave  
 CITY: Rahway  
 STATE: NJ  
 COUNTRY: USA  
 ZIP: 07065-0900  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/077,675A  
 FILING DATE: 3-JUN-1998  
 CLASSIFICATION:  
 PRIOR APPLICATION NUMBER:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Cocuzzo, Anna L.  
 REGISTRATION NUMBER: 42,452  
 REFERENCE/DOCKET NUMBER: 19590P  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 732-594-1273  
 TELEFAX: 732-594-4720  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 366 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein





Db 45 VTATCAVLFVVGISGNTLTMVLVSRFRLRTTNLYLSMAFSDLLFLC-MPLDLVRLW 103  
QY 105 RNYEPLFGPGVCYKRTALFETVCFASITITVSEVERVALHPFRALOSTRRALRL 164  
Db 104 QYRPNNEDDLCKLKFQVYSECTTATVLTALSVRFALCFPRAKVVTGKRVKVI 163  
QY 165 GIWGFVSFLPNTSINGIKFHYFPNGS-----LVPGSATCTVIRKPMIYNF 212  
Db 164 LVIMAVAFCSAGPLFVLVGEHE---NOTDPRDINECRATEFAVRSGLLIYM-VWV--- 215  
QY 213 IIQVTSFLFLPMTVLSVLYYLMALRY 240  
Db 216 -----SSVFEFLPVCELTLYLSLGRKL 238

RESULT 11  
US-09-077-675A-5  
; Sequence 5, Application US/09077675A  
; Patent No. 6242199  
; GENERAL INFORMATION:  
; APPLICANT: Pal, Lee-Yuh  
; APPLICANT: Felghner, Scott C.  
; APPLICANT: Howard, Andrew D.  
; APPLICANT: Pong, Sheng-Shung  
; APPLICANT: Van Der Ploeg, Leonardus H.T.  
; TITLE OF INVENTION: RECEPTOR ASSAY  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merck & Co., Inc.  
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
; CITY: Rahway  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065-0900  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/077,675A  
; FILING DATE: 3-JUN-1998  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cocuzzo, Anna L.  
; REGISTRATION NUMBER: 42,452  
; REFERENCE/DOCKET NUMBER: 19590P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 732-594-1273  
; TELEFAX: 732-594-4720  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 289 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-077-675A-5

Query Match 21.6%; Score 273; DB 4; Length 289;  
Best Local Similarity 31.7%; Pred. No. 1,1e-17;  
Matches 66; Conservative 47; Mismatches 69; Indels 26; Gaps 5;

QY 45 VSVVVPIRVVGVIGVAVLVLQAAKTPNNYLFSLAVSDLLVLLGMPLEVYEMW 104  
Db 46 VVATCAVLFVVGISGNTLTMVLVSRFRLRTTNLYLSMAFSDLLFLC-MPLDLVRLW 104  
QY 105 RNYEPLFGPGVCYKRTALFETVCFASITITVSEVERVALHPFRALOSTRRALRL 164

Db 105 QYRPNNEDDLCKLKFQVYSECTTATVLTALSVRFALCFPRAKVVTGKRVKVI 164  
QY 165 GIWGFVSFLPNTSINGIKFHYFPNGS-----LVPGSATCTVIRKPMIYNF 212  
Db 164 LVIMAVAFCSAGPLFVLVGEHE---NOTDPRDINECRATEFAVRSGLLIYM-VWV--- 216  
QY 213 IIQVTSFLFLPMTVLSVLYYLMALRY 240  
Db 217 -----SSVFEFLPVCELTLYLSLGRKL 239

RESULT 12  
US-07-629-1041-3  
; Sequence 3, Application US/076291041  
; Patent No. 5286621  
; GENERAL INFORMATION:  
; APPLICANT: Gershengorn, Marvin C  
; APPLICANT: Straub, Richard E  
; TITLE OF INVENTION: PITUITARY TRH RECEPTOR  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Yahwak & Associates  
; STREET: 25 Skytop Drive  
; CITY: Trumbull  
; STATE: Connecticut  
; COUNTRY: USA  
; ZIP: 06611  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Microsoft Word 3.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/629,1041  
; FILING DATE: 19901214  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: George M. Yahwak  
; REGISTRATION NUMBER: 26,824  
; REFERENCE/DOCKET NUMBER: CRF D - 995  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (203)268-1951  
; TELEFAX: (203)268-1951  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 393 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-07-629-1041-3

Query Match 21.0%; Score 265; DB 1; Length 393;  
Best Local Similarity 28.9%; Pred. No. 8,1e-17;  
Matches 72; Conservative 48; Mismatches 69; Indels 60; Gaps 9;

QY 15 LEDPFQKHINSTE-----EY-----LAFGGRSHFFLPVSVVYPIRVGVVI 58  
Db 1 MENDIVSEMNQTELOPQAAVLEIYQVITLAVIICG-----LGIIV 41  
QY 59 GAVILCVLILQAAKTPNNYLFSLAVSDLLVLLGMPLEVYEMW 114  
Db 42 GNIMVAVLVMPRTKHMRTPTNCYLSLAVADLVLAAGLPNTDSIYGSW-----VYGYV 96  
QY 115 GGYFTALEFYVCFASITITVSEVERVALHPFRALOSTRRALRLILGWGSVLF 174  
Db 97 GCLCTVLYQYLGAINSSGSIATFTIERYIAICHPITKAOFCTFSRAKKIIFVMAFTSIY 156  
QY 175 SLPNTSINGIKFHYFPNGSLVPGSATC-----TVIKPMIYNFIQVTSFLFLPMTV 228  
Db 157 CMLMFFLLDINISTYKNAVW-----SCGYKISRNYSPITLMDGV-----FTVVPML 206  
QY 229 ISVLYYMA 237

DB 207 ATVLGFI 215

## RESULT 13

US-09-261-599B-3

Sequence 3, Application US/09261599B  
Patent No. 6395877

GENERAL INFORMATION:

APPLICANT: Glucksmann, Maria A.

TITLE OF INVENTION: 14273 Receptor, A No. 6395877el G-Protein Coupled Receptor

FILE REFERENCE: 5800/4B, 03800/177086

CURRENT APPLICATION NUMBER: US/09/261,599B

PRIOR FILING DATE: 1999-02-26

PRIOR APPLICATION NUMBER: 09/107,761

PRIOR FILING DATE: 1998-06-30

PRIOR APPLICATION NUMBER: 09/223,538

PRIOR FILING DATE: 1998-12-30

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO: 3

LENGTH: 259

TYPE: PRT

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: Description of Unknown Organism: Seven

OTHER INFORMATION: Transmembrane Segment Rhodopsin Superfamily

US-09-261-599B-3

Query Match 20.8%; Score 262.5; DB 4; Length 259;

Best Local Similarity 35.3%; Pred. No. 8.7e-17;

Matches 67; Conservative 38; Mismatches 64; Indels 21; Gaps 5;

QY 59 GNVLCVLIHQAMKPTNNYLFSLVSDLVLLGMPLEVY-----EMMRNPFPLG 112

DB 1 GNLVLVLIIRKTKLPNIFLNLAVDLFLTLPPALYYLVGSEDP-----FG 55

QY 113 PGCYKTLAFTVCFASLITTSVVERVAILHPRAKLOSTR-RRALRIIGIWGFS 171

DB 56 SALCKLVTAADVNMVYASILLTAISIDRYLAIVHPLRYRRRTSPRAKVVILLVWVLA 115

QY 172 VFSLPNTSIHGKHFYFPGSLVPSACTVIKP-----MWYNIQVTSFLFY 223

DB 116 LLSLPPLLFSWKTVEEGNTLVNVVCLLDPEPESTASVTLWRSYL-LSTLVGFL 174

QY 224 LPMTVISVLY 233

DB 175 LPLVLVLCY 184

## RESULT 14

US-09-456-455A-3

Sequence 3, Application US/09456455A  
Patent No. 6448005

GENERAL INFORMATION:

APPLICANT: Glucksmann, Maria A.

TITLE OF INVENTION: 14273 Receptor, A No. 6448005el G-Protein Coupled Receptor

FILE REFERENCE: MNT-204CP3

CURRENT APPLICATION NUMBER: US/09/456,455A

PRIOR FILING DATE: 1999-12-08

PRIOR APPLICATION NUMBER: 09/107,761

PRIOR FILING DATE: 1998-06-30

PRIOR APPLICATION NUMBER: 09/223,538

PRIOR FILING DATE: 1998-12-30

NUMBER OF SEQ ID NOS: 15

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO: 3

LENGTH: 259

TYPE: PRT

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: Description of Unknown Organism: Seven

OTHER INFORMATION: Transmembrane Segment Rhodopsin Superfamily

US-09-456-455A-3

## Query Match

Best Local Similarity 35.3%; Score 262.5; DB 4; Length 259;

Matches 67; Conservative 38; Mismatches 64; Indels 21; Gaps 5;

QY 59 GNVLCVLIHQAMKPTNNYLFSLVSDLVLLGMPLEVY-----EMMRNPFPLG 112

DB 1 GNLVLVLIIRKTKLPNIFLNLAVDLFLTLPPALYYLVGSEDP-----FG 55

QY 113 PGCYKTLAFTVCFASLITTSVVERVAILHPRAKLOSTR-RRALRIIGIWGFS 171

DB 56 SALCKLVTAADVNMVYASILLTAISIDRYLAIVHPLRYRRRTSPRAKVVILLVWVLA 115

QY 172 VFSLPNTSIHGKHFYFPGSLVPSACTVIKP-----MWYNIQVTSFLFY 223

DB 116 LLSLPPLLFSWKTVEEGNTLVNVVCLLDPEPESTASVTLWRSYL-LSTLVGFL 174

QY 224 LPMTVISVLY 233

DB 175 LPLVLVLCY 184

## RESULT 15

US-08-288-663A-1

Sequence 1, Application US/08288663A  
Patent No. 5879896

GENERAL INFORMATION:

APPLICANT: HINUMA, Shuji

APPLICANT: HOSOYA, Masaki

APPLICANT: ONDA, Haruo

TITLE OF INVENTION: HUMAN TRH RECEPTOR, ITS PRODUCTION

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESS: DIKE, BRONSTEIN, ROBERTS &amp; CUSMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/288,663A

FILING DATE: 09-AUG-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 198309/1993

FILING DATE: 10-AUG-1993

APPLICATION NUMBER: 286986/1993

FILING DATE: 16-NOV-1993

APPLICATION NUMBER: 325215/1993

FILING DATE: 22-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Resnick, David S

REGISTRATION NUMBER: 34,235

REFERENCE/DOCKET NUMBER: 44612

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 398 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-288-663A-1



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: January 17, 2003, 02:26:05 : Search time 54 Seconds  
(without alignments)  
4140.136 Million cell updates/sec

Title: US-09-684-725-1

Perfect score: 729  
Sequence: 1 atggaataactcagaatgc.....cactcagagtgtatctag 729

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Indexed: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/1/lna/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/lna/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/lna/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/lna/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/lna/PCUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/lna/Backfillseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	717.8	98.5	1248	US-09-545-944-1	Sequence 1, Appl
2	282.4	38.7	1535	US-09-668-680-12	Sequence 12, Appl
3	120.4	16.5	1088	US-09-077-675A-6	Sequence 6, Appl
4	120.4	16.5	1122	US-09-077-675A-9	Sequence 9, Appl
5	114	15.6	1063	US-09-077-675A-1	Sequence 1, Appl
6	112.4	15.4	1029	US-09-077-675A-4	Sequence 4, Appl
7	112.4	15.4	1161	US-08-086-439C-2	Sequence 2, Appl
8	112.4	15.4	1161	US-08-434-877-2	Sequence 2, Appl
9	112.4	15.4	1367	US-08-475-742-3	Sequence 3, Appl
10	112.4	15.4	1370	US-08-056-051-1	Sequence 1, Appl
11	112.4	15.4	1370	US-07-928-611-17	Sequence 17, Appl
12	112.4	15.4	1370	US-08-487-811A-17	Sequence 17, Appl
13	112.4	15.4	1370	US-09-060-694-17	Sequence 17, Appl
14	112.4	15.4	1370	US-09-378-074-17	Sequence 17, Appl
15	112.4	15.4	1370	US-09-378-074-17	Sequence 17, Appl
16	112.4	15.4	1466	US-08-056-051-3	Sequence 3, Appl
17	112.4	15.4	1466	US-07-928-611-19	Sequence 19, Appl
18	112.4	15.4	1466	US-08-487-811A-19	Sequence 19, Appl
19	112.4	15.4	1466	US-09-060-694-19	Sequence 19, Appl
20	112.4	15.4	1466	US-09-378-074-19	Sequence 19, Appl
21	112.4	15.4	1466	US-09-378-074-19	Sequence 19, Appl
22	112.4	15.4	1610	US-08-056-051-5	Sequence 5, Appl
23	112.4	15.4	1610	US-07-928-611-21	Sequence 21, Appl
24	112.4	15.4	1610	US-08-487-811A-21	Sequence 21, Appl
25	112.4	15.4	1610	US-09-060-694-21	Sequence 21, Appl
26	112.4	15.4	1610	US-09-378-074-21	Sequence 21, Appl
27	112.4	15.4	1610	US-09-378-074-21	Sequence 21, Appl

28	110	15.1	1529	US-08-858-876A-3	Sequence 3, Appl
29	110	15.1	1529	US-09-472-880-3	Sequence 3, Appl
30	109.2	15.0	1092	US-09-077-675A-15	Sequence 15, Appl
31	109.2	15.0	3129	US-09-077-675A-14	Sequence 14, Appl
32	99.4	13.6	1600	US-09-029-027B-1	Sequence 1, Appl
33	99.4	13.6	1233	US-09-200-090-1	Sequence 1, Appl
34	97.6	13.4	1575	US-08-858-876A-1	Sequence 1, Appl
35	97.6	13.4	1575	US-09-472-880-1	Sequence 1, Appl
36	97.2	13.3	283	US-08-993-088A-4	Sequence 4, Appl
37	97.2	13.3	283	US-08-993-424B-4	Sequence 4, Appl
38	95.8	13.1	1176	US-09-200-090-3	Sequence 3, Appl
39	93.2	12.8	1164	US-08-993-088A-6	Sequence 6, Appl
40	93.2	12.8	1164	US-08-993-424B-6	Sequence 6, Appl
41	88.8	12.2	1601	US-08-722-001-7	Sequence 7, Appl
42	88.8	12.2	1997	US-08-722-001-27	Sequence 27, Appl
43	88.8	12.2	2004	US-08-722-001-11	Sequence 11, Appl
44	88.6	12.2	1639	US-08-334-698-5	Sequence 5, Appl
45	88.6	12.2	1639	US-08-228-932-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-09-545-944-1  
; Sequence 1, Application US/09545944  
; Patent No. 6461836  
; GENERAL INFORMATION:  
; APPLICANT: AMES, ROBERT  
; APPLICANT: ELSHOURBAGY, NABIL  
; APPLICANT: MICHALOVICH, DAVID  
; APPLICANT: SARAU, HENRY  
; APPLICANT: SHARON, USMAN  
; APPLICANT: VAMTER, LISA  
; TITLE OF INVENTION: MOLECULAR CLONING OF A 7TM RECEPTOR  
; FILE REFERENCE: (AXOR34) AND SCREENING METHODS THEREOF  
; CURRENT APPLICATION NUMBER: US/09/545,944  
; CURRENT FILING DATE: 2000-04-10  
; PRIOR APPLICATION NUMBER: US 09/435,384  
; PRIOR FILING DATE: 1999-11-05  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1248  
; TYPE: DNA  
; ORGANISM: HOMO SAPIENS  
; US-09-545-944-1

Query Match 98.5%: Score 717.8; DB 4; Length 1248;  
Best Local Similarity 99.7%: Pred. No. 1.4e+149;  
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAAACCTTCAGATCTCTCTGATCTACACAGAACTAGAGATCTTCAG 60  
|||||  
DB 10 ATGGAAACCTTCAGATCTCTCTGATCTACACAGAACTAGAGATCTTCAG 69  
|||||  
QY 61 AAACACCTGACGACCGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
|||||  
DB 70 AAACACCTGACGACCGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 129  
|||||  
QY 121 TCTTCT 180  
|||||  
DB 130 TCTTCT 189  
|||||  
QY 181 GTCCTGCTGCT 240  
|||||  
DB 190 GTCCTGCTGCT 249  
|||||  
QY 241 CTCCTTACGCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
|||||  
DB 250 CTCCTTACGCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 309  
|||||



```

: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Cocuzzo, Anna L.
: REGISTRATION NUMBER: 42,452
: REFERENCE/DOCKET NUMBER: 19590P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 732-594-1273
: TELEFAX: 732-594-4720
: TELEX:
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1088 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
US-09-077-675A-6

```

```

Query Match      16.5%; Score 120.4; DB 4; Length 1088;
Best Local Similarity 52.3%; Pred. No. 4.2e-18;
Matches 290; Conservative 0; Mismatches 261; Indels 3; Gaps 1;

```

```

QY 132 CGTGTCTGTGCTATGATGCCAATTTTGTGTGGGGGTCATTGGCAATGTCCTGGTGTG 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 CGTCAAGCCACCTGGCTGCTCTGTGTGGTATGCTGCAACCTGCTCACCAT 181
QY 192 CCGTGATTTCTGACAGCAGCCTATGAAGAGCCGACCACTACTACTCTTCAAGCCT 251
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 GGTGTGTGTGTCGCTTCCGAGCTGCGCACACCACCACTTCACTGTCACAGCAT 241
QY 252 GCGGCTCTGACCTCTGCTGCTCTTGGAAATGCCCTGGAGTCTATGATGATGTG 311
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 GGCCTTCTCCGAT---CTGCTCATCTTCTCTGATATGCCCTGAGCTCTTGGCCTGTG 238
QY 312 GCCCAACTACCTCTTCTTGTGTGGGCCGCTGGGCTGCTACTTCAAGAGGCCCTTTGA 371
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 GCAGTACCGGCTCGGAACCTTCGGCGACCTCTCTGCAAACTCTTCAATTCGTCAAGTGA 358
QY 372 GACCGTGTGCTTGGCCTCCATCTCTCAGCATCACCACCGTACGGTGAAGGCTACGAGGC 431
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 GACCTGACCTACCGCAGGCTGCTCAGCATCAGCAGCGCTAGGCTGACAGGCTACTCCGC 418
QY 432 CATCTACACCCGCTCCGCGCCAAACTGACAGACCCCGCGCCGCTCAGATCCT 491
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 419 CATCTGCTTCCACTCCGGGCCAAGGTGTGTGATCACCAGGGGGGTGAAGCTGGTCAT 478
QY 492 CGGCATGCTGTGGGGCTTCTCGTGTCTTCTCTGCTGCCCCAACAACAGCATTCGAT 551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 479 CTTCGTCATCTGGGCGGTGGCTTCTGTGACAGCGCGGGCCCATCTTGTGCTAGTGGGGT 538
QY 552 CAAGTTCACACTTCTCCCAATGGTCCCTGTGTCACAGTTCGGCCACCTGTACGGTCAT 611
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 539 GAGACAGCAAGAGGACCGACCTTGGGACCAACAGATGCGCGCCACCGAGTTTGC 598
QY 612 CAAGCCATGTGATCTACATTTTCATCATCAGATCACCCTCTCTATTACCTCCT 671
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 599 GGTGGCTGTGAGCTGCTCAGCGTCAATGATGTGGGTGTCCAGCATCTTCTTCTCTCC 658
QY 672 CCCCAGTACTGCA 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 659 TGTCTTCTGTCTCA 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 4
US-09-077-675A-9
: Sequence 9, Application US/09077675A
: Patent No. 6242199
: GENERAL INFORMATION:
: APPLICANT: Pal, Lee-Yuh
: APPLICANT: Feighner, Scott C.
: APPLICANT: Howard, Andrew D.
: APPLICANT: Pong, Sheng-Shung
: APPLICANT: Van Der Ploeg, Leonardus H.T.

```

```

: TITLE OF INVENTION: RECEPTOR ASSAY
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merck & Co., Inc.
: STREET: P.O. Box 2000, 126 E. Lincoln Ave.
: CITY: Rahway
: STATE: NJ
: COUNTRY: USA
: ZIP: 07065-0900
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/077,675A
: FILING DATE: 3-JUN-1998
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Cocuzzo, Anna L.
: REGISTRATION NUMBER: 42,452
: REFERENCE/DOCKET NUMBER: 19590P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 732-594-1273
: TELEFAX: 732-594-4720
: TELEX:
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1122 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
US-09-077-675A-9

```

```

Query Match      16.5%; Score 120.4; DB 4; Length 1122;
Best Local Similarity 52.3%; Pred. No. 4.2e-18;
Matches 290; Conservative 0; Mismatches 261; Indels 3; Gaps 1;

```

```

QY 132 CGTGTCTGTGCTATGATGCCAATTTTGTGTGGGGGTCATTGGCAATGTCCTGGTGTG 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 387 GGTACACAGCCACCTGCTGTGCTCTTCTGTGTGATGCTGTGCACTGCTCACCAT 446
QY 192 CCGTGATTTCTGACAGCAGCCTATGAAGAGCCGACCACTACTACTCTTCAAGCCT 251
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 447 GGTGTGTGTGCGGCTTCCGCGAGGTGGCGACACACCACTTCACTGTCAGCAT 506
QY 252 GCGGCTCTGACCTCTGCTGCTCTTGGAAATGCCCTGGAGTCTATGATGATGTG 311
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 507 GGCCTTCTCCGAT---CTGCTCATCTTCTCTGATATGCCCTGAGACTCGTGGCCTGTG 563
QY 312 GCGCACTACCTCTTCTTGTGTGGGCCGCTGGGCTGCTACTTCAAGAGGCCCTTTGA 371
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 564 GCAGTACCGGCTCGAATCTTGCGACCTCTCTGTGAAACTCTTCAATTCGTCAAGTGA 623
QY 372 GACCGTGTGCTGCTCTCATCTCTCAGCATCACCACCGTACAGGTGAGCGCTACGTGCG 431
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 624 GAGCTGACCTTACGCAAGGCTGCTCAGCATCAGCGCTGAGGCTGAGGCTACTTCGC 683
QY 432 CATCTACACCCGCTCCGCGCCAAACTGACAGAGACCCGCGCGCGCTCAGATCCT 491
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 684 CATCTGCTTCCACTCCGGGCCAAGGTGTGTGATCACCAGAGGCGGTGAAGCTGTGTCAT 743
QY 492 CCGCATGCTGTGGGGCTTCCGATCTTCTCTGCTGCTGCAACACAGCATTCATGTCAT 551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 744 CTTCGTCATCTGGGCGCTGCTTCTGTGACAGCGCGGGCCCATCTTGTGCTAGTGGGGT 803
QY 552 CAAGTTCACACTTCTCCCAATGGTCCCTGTGTCACAGTTCGGCCACCTGTACGGTCAT 611
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 804 GAGACAGCAAGAGGACCGACCTTGGAGACACCAAGATGTGCGGCCACCGAGTTTGC 863

```







```
STATE: Michigan
COUNTRY: USA
ZIP: 49001
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette (DS,HD)
COMPUTER: Gateway 2000, P5-90
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,877
FILING DATE: 1 July 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Darinley Jr., James D.
REGISTRATION NUMBER: 33,673
REFERENCE/DOCKET NUMBER: 4700 DVI
TELECOMMUNICATION INFORMATION:
TELEPHONE: 616-385-5210
TELEFAX: 616-385-6897
TELEX: 224401
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1161 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-434-877-2
```

```
Query Match          15.4%; Score 112.4; DB 1; Length 1161;
Best Local Similarity 53.4%; Pred. No. 2.5e-16;
Matches 236; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 139 GTGGGTATGTGCAATTTTGTGGGGGTCATGTGCAATGTCTGTGTCTGTG 198
DB 112 GTGGGGGGGCTGTCTCATCGCGGGTCTGCGGGAACCTGCTGCTGCGTGA 171
QY 199 ATTGTGACACAGGAGCTTGAAGAGCCGACCACTACTCTTTCAGCCGTGGG 258
DB 172 GTGGGACAGGAGGCGGCGTGCAGAGCCGACCACTCTTCATGCTGAGCCGTG 231
QY 259 TGTGACTCTGTGCTGTCTCTTGAATGCCCTGGAGGCTAAGATGTGGCGCA 318
DB 232 GCGGACCTCTCTCTGCTCTGTGGTGGCGGCTCTTCTTACTCCGAGGTCCAG 291
QY 319 TACCTTTCTTGTGGGGCGGCGGAGCTGTCTTCAAGACGCGCTTTTGAAGC 378
DB 292 GCGGCTGTGCTGTAGACCCCGCTGTGCAAGCCCTCATGAGCATGAGCATG 351
379 TGCCTTCCGCTCATCTTCAGATCAACACGCTGAGGCTGAGCGGTACGTGCA 438
DB 352 TGCACGCGCTCATCTTCAACCTGTGCGGCATCAGGTGACAGTTCGTGGCCG 411
QY 439 CACCGGTTCCGGCCAACTGACAGACACCGCGCGCGGCGGCTCAGATCTCG 498
DB 412 GTGCGCTGTGCTTACAAACGCGAGGTGAGACCGCGGCACTGCTCATGCG 471
QY 499 GTCTGGGGCTTCTCGGTCTTTCCTCCGCAACACAGATCCATGAGCATCA 558
DB 472 ACGTGGCTGTCTCGCGCGGCTGGGCGGCGGCTTACTGTGTGGGCTCAAC 531
QY 559 CACTACTTCCCAATGGGTCC 580
DB 532 GCGCGGACCGCGCGGTGTGC 553
```

```
RESULT 9
US-08-475-742-3
; Sequence 3, Application US/08475742
; Patent No. 6121015
; GENERAL INFORMATION:
; APPLICANT: O'Malley, Karen L
; APPLICANT: Todd, Richard D
```

```
TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor
FILE REFERENCE: WU 102 CON DIV
CURRENT APPLICATION NUMBER: US/08/475,742
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: US 08/261,293
EARLIER FILING DATE: 1994-06-16
EARLIER APPLICATION NUMBER: US 08/014,013
EARLIER FILING DATE: 1993-01-28
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 1367
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)..(1367)
OTHER INFORMATION: D4 Dopamine Receptor cDNA
PUBLICATION INFORMATION:
AUTHORS: Van Tol, H. H.
AUTHORS: Bunzow, J. R.
TITLE: Cloning of the gene for a human dopamine D4 receptor
TITLE: with high affinity for the antipsychotic clozapine
JOURNAL: Nature
VOLUME: 350
PAGES: 610-614
DATE: 1991
US-08-475-742-3
```

```
Query Match          15.4%; Score 112.4; DB 3; Length 1367;
Best Local Similarity 53.4%; Pred. No. 2.5e-16;
Matches 236; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 139 GTGGGTATGTGCAATTTTGTGGGGGTCATGTGCAATGTCTGTGTCTGTG 198
DB 218 GTGGGGGGGCTGTCTCATCGCGGGTCTGCGGGAACCTGCTGTGCTGCA 277
QY 199 ATTGTGACACAGGAGCTTGAAGAGCCGACCACTACTCTTTCAGCCGTGGG 258
DB 278 GTGGGACAGGAGGCGGCGTGCAGAGCCGACCACTCTTCTCATGCTGAGCC 337
QY 259 TGTGACCTCTGTGCTGTCTCTTGAATGCCCTGAGAGTATGAGATGGCGCA 318
DB 338 GCGGACCTCTCTGCTGTCTGTGTGCTGCTGCTCTTCTTACTCCGAGGTCC 397
QY 319 TACCTTTCTTGTGGGGCGGCGGAGCTGTCTTCAAGACGCGCTTTTGAAGC 378
DB 398 GCGGCTGTGCTGTAGACCCCGCTGTGCAAGCCCTCATGAGCATGAGCATG 457
QY 379 TGCCTTCCGCTCATCTTCAGATCAACACGCTGAGGCTGAGCGGTACGTGCA 438
DB 458 TGCACGCGCTCATCTTCAACCTGTGCGGCATCAGCGTGTGAGGCTGTGCG 517
QY 439 CACCGGTTCCGGCCAACTGACAGACACCGGCGCGGCGGCTCAGATCTCG 498
DB 518 GTGCGGCTGTGCTTCAACCGGAGGCTGGAGCGCGCGGCACTCTCTATCG 577
QY 499 GTCTGGGGCTTCTCGGTCTTTCCTCCGCAACACAGATCCATGAGCATCA 558
DB 578 ACGTGGCTGTCTCGCGCGGCTGGGCGGCGGCTTACTGTGTGGGCTCAAC 637
QY 559 CACTACTTCCCAATGGGTCC 580
DB 638 GCGCGGACCGCGCGGTGTGC 659
```

```
RESULT 10
US-08-056-051-1
; Sequence 1, Application US/08056051
; Patent No. 5516683
; GENERAL INFORMATION:
; APPLICANT: Grandy, David K
; APPLICANT: Bunzow, James R
```







```

RESULT 15
PCT-US93-07370-17
: Sequence 17, Application PC/TUS93/07370
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses
: NUMBER OF SEQUENCES: 22
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/07370
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1370 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: 1..103
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: 1268..1370
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 104..1267
: PCT-US93-07370-17

```

[illegible]

Search completed: January 17, 2003, 02:35:14  
Job time : 61 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 02:36:41 ; Search time 72 Seconds  
(without alignments)  
447.870 Million cell updates/sec

Title: US-09-684-725-2  
Perfect score: 1263  
Sequence: 1 MEKLNASWYQOKLEDPFO.....LLPMTVISLVYLMALRVSI 242

oring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_101002:\*

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1263	100.0	242	22	AA668333
2	1252	99.1	249	22	AA680937
3	1252	99.1	293	22	AA667807
4	1252	99.1	296	22	AA667806
5	1252	99.1	412	22	AA667803
6	1252	99.1	412	22	AA667805
7	1252	99.1	415	21	AA771296
8	1252	99.1	415	21	AA602830
9	1252	99.1	415	21	AA552992
10	1252	99.1	415	22	AA663353

11	1252	99.1	415	22	AA663366	Amino acid sequenc
12	1252	99.1	415	22	AA664297	Human GTP-binding
13	1252	99.1	415	22	AA603629	Human G-protein co
14	1252	99.1	415	22	AA667802	Amino acid sequenc
15	1252	99.1	415	22	AA667804	Amino acid sequenc
16	1252	99.1	415	23	AA077155	Human G-protein co
17	1252	99.1	415	23	AAE14262	Human NMUR2 protei
18	1033.5	81.8	395	22	AA663367	Amino acid sequenc
19	1033.5	81.8	395	22	AAE03634	Rat G-protein coup
20	1033.5	81.8	395	23	AAE14263	Rat NMUR2 protein.
21	680.5	53.9	403	21	AAV90638	Human G protein-co
22	680.5	53.9	403	21	AAV90673	Human mutant G pro
23	680.5	53.9	403	21	AAV44642	Human growth hormo
24	680.5	53.9	403	22	AA699185	Human FM-3. Homo
25	680.5	53.9	403	22	AAE03628	Human G-protein co
26	677.5	53.6	445	22	AA068523	Human novel cytol
27	677.5	53.6	445	22	AA068566	Human novel cytol
28	671.5	53.2	426	23	AAE15631	Human G-protein co
29	646	51.1	413	22	AAE03635	Rat G-protein coup
30	642	50.8	412	22	AA699199	Rat FM-3. Rattus
31	642	50.8	439	22	AAE03636	Rat G-protein coup
32	586	47.2	405	21	AAV44641	Mouse growth hormo
33	586	47.2	405	22	AA699186	Murine FM-3. Mus
34	418	33.1	419	22	AA686964	D. melanogaster pe
35	418	33.1	428	22	ABB64752	Drosophila melanog
36	418	33.1	428	22	AAU38960	Drosophila G-prote
37	352	27.9	595	22	ABB71697	Drosophila melanog
38	352	27.9	595	22	AAU38987	Drosophila G-prote
39	337	26.7	660	22	AA671689	Drosophila melanog
40	337	26.7	660	22	AAU38986	Drosophila G-prote
41	333	26.4	386	21	AAV54146	Amino acid sequenc
42	333	26.4	386	22	AA662653	Short form of moti
43	333	26.4	412	21	AA602854	Human G protein co
44	333	26.4	412	21	AAV54145	Amino acid sequenc
45	333	26.4	412	22	AA662652	Long form of motil

ALIGNMENTS

RESULT 1

AA668333

ID AAB68333 standard; Protein; 242 AA.

XX

AC AAB68333;

XX

DT 09-JUL-2001 (first entry)

XX

DE Amino acid sequence of human g-protein coupled receptor PFT-002.

XX

KW G-protein coupled receptor; obesity; signal transduction; diabetes;

KW metabolic disease; neurological disease; psychotherapy; dermatology;

KW urogenital disease; inflammation; cancer; tissue repair; photoaging;

KW skin pigmentation; frailty; osteoporosis; cardiovascular disease;

KW gastrointestinal disease; infection; allergy; respiratory disease;

KW sensory organ disorder; sleep disorder; hair loss; gene therapy;

KW PFT-002.

XX

OS Homo sapiens.

XX

PN EP1090990-A1.

XX

PD 11-APR-2001.

XX

PD 06-OCT-2000; 2000EP-0308852.

XX

PR 08-OCT-1999; 99GB-0023888.

XX

PA (PFT2 ) PFIZER LTD.

PA (PFT2 ) PFIZER INC.

XX

PA Harland L;

XX

DR WPI: 2001-302046/32.  
 DR N-PSDB; AAF85107.  
 XX  
 PT New human G-protein coupled receptor (GPCR) polynucleotides and  
 PT polypeptides, for screening modulators of the polypeptide useful in  
 PT treating diseases associated with signal transduction, e.g. cancer,  
 PT inflammation, or especially, obesity -  
 XX  
 PS Claim 22: Page 42-43; 53pp; English.  
 XX  
 CC The present sequence represents a human G-protein coupled receptor. The  
 CC G-protein coupled receptor polynucleotide and polypeptide are useful  
 CC as pharmaceuticals or in the manufacture of medicaments for the  
 CC treatment of obesity. They are useful in the diagnosis and treatment  
 CC of diseases and disorders associated with signal transduction such as  
 CC obesity, diabetes and metabolic disease, neurological disease, cancer,  
 CC psychopharmacology, urogenital disease, inflammation, cancer, tissue  
 CC repair, dermatology, skin pigmentation, photoregulation, frailty,  
 CC osteoporosis, cardiovascular disease, gastrointestinal disease,  
 CC infection, allergy and respiratory disease, sensory organ disorders,  
 CC sleep disorders and hair loss. The polynucleotide may also be useful  
 CC in gene therapy.  
 CC  
 XX  
 SQ Sequence 242 AA:  
 Query Match 100.0%; Score 1263; DB 22; Length 242;  
 Best Local Similarity 100.0%; Pred. No. 1e-142;  
 Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MERLQNASWITVOOKLEDPFOKHLNSTEETLAFICGPRRSHFPLPSVYVPIFVGVIGN 60  
 PT |  
 PT |  
 PT 1 MERLQNASWITVOOKLEDPFOKHLNSTEETLAFICGPRRSHFPLPSVYVPIFVGVIGN 60  
 DB |  
 QY 61 VLVCIVIIHQHAKPTPNNTYLFSLAVSDLLVLLGMPLEVEEMRNPFEGVCYFKT 120  
 DB 61 VLVCIVIIHQHAKPTPNNTYLFSLAVSDLLVLLGMPLEVEEMRNPFEGVCYFKT 120  
 QY 121 ALFEIVCFASILSTTVSVERVYAILHPFRAKLOSTRRALRIIGIWGFSVLSLPTS 180  
 DB 121 ALFEIVCFASILSTTVSVERVYAILHPFRAKLOSTRRALRIIGIWGFSVLSLPTS 180  
 QY 181 IHGIKHYFPNGSLVPGSACCTYIKPMWITNFIQVTSFLFLLPMIVISLVLYMALRV 240  
 DB 181 IHGIKHYFPNGSLVPGSACCTYIKPMWITNFIQVTSFLFLLPMIVISLVLYMALRV 240  
 QY 241 SI 242  
 DB 241 SI 242  
 RESULT 2  
 AAG80937  
 ID AAG80937 standard; Protein; 249 AA.  
 AAG80937;  
 XX 28-AUG-2001 (first entry)  
 XX  
 XX Human nGPR15.  
 DE  
 XX  
 XX G protein-coupled receptor; nGPR: seven transmembrane receptor;  
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;  
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;  
 KW cardiovascular disease; proliferative disorder; hormonal disorder;  
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;  
 KW attention deficit-hyperactivity disorder/attention deficit disorder;  
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;  
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;  
 KW neuroprotective.  
 KW  
 XX Homo sapiens.  
 OS  
 XX  
 PN MO200136473-A2.

XX  
 PD 25-MAY-2001.  
 XX  
 XX 16-NOV-2000; 2000MO-US31581.  
 PE 16-NOV-1999; 99US-0165838.  
 PR 17-NOV-1999; 99US-0166071.  
 PR 19-NOV-1999; 99US-0166678.  
 PR 28-DEC-1999; 99US-0173396.  
 PR 22-FEB-2000; 2000US-0184129.  
 PR 28-FEB-2000; 2000US-0185421.  
 PR 02-MAR-2000; 2000US-0185554.  
 PR 02-MAR-2000; 2000US-0186530.  
 PR 03-MAR-2000; 2000US-0186811.  
 PR 09-MAR-2000; 2000US-0188114.  
 PR 17-MAR-2000; 2000US-0190310.  
 PR 21-MAR-2000; 2000US-0190800.  
 PR 20-APR-2000; 2000US-0198568.  
 PR 02-MAY-2000; 2000US-0201190.  
 PR 08-MAY-2000; 2000US-0203111.  
 PR 25-MAY-2000; 2000US-0207094.  
 XX  
 PA (PHMA ) PHARMACIA & UPJOHN CO.  
 XX  
 XX  
 PI Vogeli G, Wood LS, Parodi LA, Hiesch RR, Lind P, Slightom J;  
 PI Schellin KA, Kayes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;  
 XX  
 DR WPI: 2001-389826/41.  
 DR N-PSDB; AAH50977.  
 XX  
 PT New G protein-coupled receptor (nGPR-x) and its encoding  
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -  
 XX  
 PS Claim 37: Page 79; 261pp; English.  
 XX  
 CC The present invention relates to novel G protein-coupled receptors  
 CC (nGPRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,  
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present  
 CC sequence is one such G protein-coupled receptor. GPCRs are also known as  
 CC seven transmembrane receptors and function in signal transduction. The  
 CC nGPRx coding sequences are useful for screening a human to diagnose a  
 CC disorder affecting the brain or a genetic predisposition, specifically  
 CC schizophrenia. nGPRx are useful for identifying compounds useful for  
 CC treating schizophrenia. Detection of nGPRx in a sample is useful as a  
 CC diagnostic tool for diseases or disorders e.g. thyroid disorders, renal  
 CC failure, rheumatoid arthritis, CNS disorders, infectious such as HIV-1,  
 CC metabolic and cardiovascular diseases, proliferative disorders and  
 CC hormonal disorders. Modulators of nGPRx activity have the utility for  
 CC treating neurological disorders, including schizophrenia, ADHD/ADD  
 CC (attention deficit-hyperactivity disorder/attention deficit disorder),  
 CC and neuronal disorders such as Alzheimer's disease, Parkinson's disease,  
 CC migraine and senile dementia. Additional disorders include inflammatory  
 CC conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune  
 CC disorders, cancers, respiratory ailments such as asthma, and inflammatory  
 CC diseases e.g. inflammatory bowel disease.  
 CC  
 XX  
 SQ Sequence 249 AA:  
 Query Match 99.1%; Score 1252; DB 22; Length 249;  
 Best Local Similarity 99.2%; Pred. No. 2.3e-141;  
 Matches 240; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MERLQNASWITVOOKLEDPFOKHLNSTEETLAFICGPRRSHFPLPSVYVPIFVGVIGN 60  
 DB 3 MERLQNASWITVOOKLEDPFOKHLNSTEETLAFICGPRRSHFPLPSVYVPIFVGVIGN 62  
 QY 61 VLVCIVIIHQHAKPTPNNTYLFSLAVSDLLVLLGMPLEVEEMRNPFEGVCYFKT 120  
 DB 63 VLVCIVIIHQHAKPTPNNTYLFSLAVSDLLVLLGMPLEVEEMRNPFEGVCYFKT 122  
 QY 121 ALFEIVCFASILSTTVSVERVYAILHPFRAKLOSTRRALRIIGIWGFSVLSLPTS 180  
 DB 123 ALFEIVCFASILSTTVSVERVYAILHPFRAKLOSTRRALRIIGIWGFSVLSLPTS 182





CC The present sequence represents a splice variant of the long version of a  
 CC human G-protein coupled receptor designated IGS4A. IGS4 exists in two  
 CC polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4  
 CC polynucleotides are useful for preventing, ameliorating or correcting  
 CC dysfunctions or diseases. These diseases include peripheral nervous  
 CC system, psychiatric and central nervous system disorders  
 CC (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's  
 CC disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases  
 CC (e.g. heart failure, angina pectoris, myocardial infarction or  
 CC hypertension), dyslipidemias, obesity, emesis, gastrointestinal  
 CC disorders (e.g. inflammatory bowel disease or motility disorders),  
 CC osteoporosis, inflammations, infections (e.g. bacterial, fungal,  
 CC protozoan or viral), pain, cancers, immune disorders, allergies,  
 CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are  
 CC effective with regard to disorders of the nervous system, including the  
 CC central and peripheral nervous systems, disorders of the gastrointestinal  
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or  
 CC genitourinary system, or immunological disease. The IGS4 polynucleotides  
 CC are useful as diagnostic reagents for detecting under-expression,  
 CC overexpression or altered expression of IGS4.

XX Sequence 296 AA;

Query Match 99.1%; Score 1252; DB 22; Length 296;  
 Best Local Similarity 99.6%; Pred. No. 2.8e-141;  
 Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKLNQASWIIYQOKLEDDPFQKHLNSTEELAFICGPRSHFFLPVSVYVPFVGVGIGN 60  
 DB 4 MEKLNQASWIIYQOKLEDDPFQKHLNSTEELAFICGPRSHFFLPVSVYVPFVGVGIGN 63  
 QY 61 VVLCVYIIHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMRMNYPFLFGVGCYFRT 120  
 DB 64 VVLCVYIIHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMRMNYPFLFGVGCYFRT 123  
 QY 121 ALFETVCFASILSITTVSERYVAIIHPFRAKLOSTRRALRLIGYWGFSVLSLPNTS 180  
 DB 124 ALFETVCFASILSITTVSERYVAIIHPFRAKLOSTRRALRLIGYWGFSVLSLPNTS 183  
 QY 181 HGIKRHFPPNSLVPGSATCTVIRKPMIYNFTIOVTSFLFLLPMTVISVLYYLMALRV 240  
 DB 184 HGIKRHFPPNSLVPGSATCTVIRKPMIYNFTIOVTSFLFLLPMTVISVLYYLMALRV 243

RESULT 5  
 AAB67803  
 ID AAB67803 standard; Protein; 412 AA.

XX AAB67803;

DT 29-JUN-2001 (first entry)

DE Amino acid sequence of G-protein coupled receptor IGS4A short version.

Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;  
 nervous system disorder; psychiatric disorder; Parkinson's disease;  
 episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;  
 cardiovascular disease; heart failure; angina pectoris; obesity; emesis;  
 motility disorder; myocardial infarction; hypertension; dyslipidemia;  
 gastrointestinal disorder; inflammatory bowel disease; osteoporosis;  
 inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;  
 gynecological disorder.

XX Homo sapiens.

OS WO200125269-A2.

XX 12-APR-2001.

PF 25-SEP-2000; 2000WO-EP09584.

XX 24-SEP-1999; 99EP-0203140.  
 PR 24-SEP-1999; 99NL-1013140.

PR 28-JUL-2000; 2000EP-0202683.  
 PR 31-JUL-2000; 2000US-0222047.  
 XX (SOLV ) SOLVAY PHARM BV.  
 PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;  
 DR WPI: 2001-273568/28.  
 XX N-PSDB; AAF80323.

PT New G-protein coupled receptors and the polynucleotides encoding them,  
 PT useful for preventing, ameliorating or correcting nervous system  
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain  
 or cancers  
 XX  
 PS Claim 18; Page 81-82; 102pp; English.

CC The present sequence represents the short version of a human G-protein  
 CC coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,  
 CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful  
 CC for preventing, ameliorating or correcting dysfunctions or diseases.  
 CC These diseases include peripheral nervous system, psychiatric and central  
 CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal  
 CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or  
 CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,  
 CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,  
 CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility  
 CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,  
 CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,  
 CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are  
 CC effective with regard to disorders of the nervous system, including the  
 CC central and peripheral nervous systems, disorders of the gastrointestinal  
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or  
 CC genitourinary system, or immunological disease. The IGS4 polynucleotides  
 CC are useful as diagnostic reagents for detecting under-expression,  
 CC overexpression or altered expression of IGS4.

Sequence 412 AA;  
 Query Match 99.1%; Score 1252; DB 22; Length 412;  
 Best Local Similarity 99.6%; Pred. No. 4.4e-141;  
 Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKLNQASWIIYQOKLEDDPFQKHLNSTEELAFICGPRSHFFLPVSVYVPFVGVGIGN 60  
 DB 1 MEKLNQASWIIYQOKLEDDPFQKHLNSTEELAFICGPRSHFFLPVSVYVPFVGVGIGN 60  
 QY 61 VVLCVYIIHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMRMNYPFLFGVGCYFRT 120  
 DB 61 VVLCVYIIHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMRMNYPFLFGVGCYFRT 120  
 QY 121 ALFETVCFASILSITTVSERYVAIIHPFRAKLOSTRRALRLIGYWGFSVLSLPNTS 180  
 DB 121 ALFETVCFASILSITTVSERYVAIIHPFRAKLOSTRRALRLIGYWGFSVLSLPNTS 180  
 QY 181 HGIKRHFPPNSLVPGSATCTVIRKPMIYNFTIOVTSFLFLLPMTVISVLYYLMALRV 240  
 DB 181 HGIKRHFPPNSLVPGSATCTVIRKPMIYNFTIOVTSFLFLLPMTVISVLYYLMALRV 240

RESULT 6  
 AAB67805  
 ID AAB67805 standard; Protein; 412 AA.

XX AAB67805;

DT 29-JUN-2001 (first entry)

DE Amino acid sequence of G-protein coupled receptor IGS4B short version.

Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;  
 nervous system disorder; psychiatric disorder; Parkinson's disease;  
 episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;

KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;  
 KW motility disorder; myocardial infarction; hypertension; dyslipidemia;  
 KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;  
 KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;  
 KW gynecological disorder.  
 OS Homo sapiens.  
 PN WO200125269-A2.  
 XX  
 XX  
 PD 12-APR-2001.  
 XX  
 XX 25-SEP-2000; 2000WO-EP09584.  
 PF 24-SEP-1999; 99EP-0203140.  
 XX 24-SEP-1999; 99NL-1013140.  
 XX 28-JUL-2000; 2000EP-0202683.  
 PR 31-JUL-2000; 2000US-0222047.  
 PA (SOLV ) SOLVAY PHARM BV.  
 PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;  
 DR WPI; 2001-273568/28.  
 DR N-PSDB; AAF80325.  
 PT New G-protein coupled receptors and the polynucleotides encoding them,  
 PT useful for preventing, ameliorating or correcting nervous system  
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain  
 PT or cancers -  
 PS  
 PS Claim 18; Page 89-90; 102pp; English.  
 XX  
 CC The present sequence represents the short version of a human G-protein  
 CC coupled receptor designated IG54B. IG54 exists in two polymorphic forms,  
 CC IG54A and IG54B. The IG54 receptors and IG54 polynucleotides are useful  
 CC for preventing, ameliorating or correcting dysfunctions or diseases.  
 CC These diseases include peripheral nervous system, psychiatric and central  
 CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal  
 CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or  
 CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,  
 CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,  
 CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility  
 CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,  
 CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,  
 CC sepsis or gynecological disorders. Agonists or antagonists of IG54 are  
 CC effective with regard to disorders of the nervous system, including the  
 CC central and peripheral nervous systems, disorders of the gastrointestinal  
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or  
 CC genitourinary system, or immunological disease. The IG54 polynucleotides  
 CC are useful as diagnostic reagents for detecting under-expression,  
 CC overexpression or altered expression of IG54.  
 CC  
 CC  
 SQ Sequence 412 AA:  
 XX  
 XX  
 Query Match 99.1%; Score 1252; DB 22; Length 412;  
 Best Local Similarity 99.6%; Pred. No. 4.4e-141;  
 Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 181 IHGKIFHPNGSLVPGSATCTVTKPMWYNTFIQVTSFLFLLPMTVISLVLYMALRL 240  
 RESULT 7  
 ID AAY71296  
 AC AAY71296 standard; Protein; 415 AA.  
 XX  
 XX AAY71296;  
 DT 02-NOV-2000 (first entry)  
 XX  
 DE Human orphan G protein-coupled receptor hrup6.  
 XX  
 KW Human; orphan G protein-coupled receptor; GPCR; hrup6; drug screening;  
 KW transmembrane receptor; signal cascade.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200031258-A2.  
 PD 02-JUN-2000.  
 XX  
 PF 13-OCT-1999; 99WO-US23687.  
 XX  
 XX 20-NOV-1998; 98US-0109213.  
 XX 16-FEB-1999; 99US-0120416.  
 PR 26-FEB-1999; 99US-0121852.  
 PR 12-MAR-1999; 99US-0123946.  
 PR 12-MAR-1999; 99US-0123949.  
 PR 28-MAY-1999; 99US-0136436.  
 PR 28-MAY-1999; 99US-0136437.  
 PR 28-MAY-1999; 99US-0136439.  
 PR 28-MAY-1999; 99US-0136567.  
 PR 28-MAY-1999; 99US-0137127.  
 PR 28-MAY-1999; 99US-0137131.  
 PR 29-JUN-1999; 99US-0141448.  
 PR 29-SEP-1999; 99US-0156555.  
 PR 29-SEP-1999; 99US-0156633.  
 PR 29-SEP-1999; 99US-0156634.  
 PR 29-SEP-1999; 99US-0156653.  
 PR 01-OCT-1999; 99US-0157280.  
 PR 01-OCT-1999; 99US-0157281.  
 PR 01-OCT-1999; 99US-0157282.  
 PR 01-OCT-1999; 99US-0157293.  
 PR 01-OCT-1999; 99US-0157294.  
 PR 12-OCT-1999; 99US-0416760.  
 PR 12-OCT-1999; 99US-0417044.  
 PA (AREN-) ARENA PHARM INC.  
 PI  
 PI Chen R, Dang HT, Liaw CW, Lin I;  
 DR WPI; 2000-400068/34.  
 DR N-PSDB; AAD01123.  
 PT Novel human orphan G protein-coupled receptors and the encoding cDNAs  
 PT for use in the identification of G protein-coupled receptor agonists -  
 PT  
 PS Claim 22; Page 57-59; 102pp; English.  
 XX  
 XX The present amino acid sequence is the hrup6, an endogenous human  
 XX orphan G protein-coupled receptor (GPCR). The full length hrup6 cDNA  
 XX was cloned by RT-PCR using human thymus cDNA as template.  
 CC The orphan GPCR of the invention, like all GPCRs has seven transmembrane  
 CC alpha helices with an extracellular N-terminus and an intracellular  
 CC C-terminus. However, no endogenous ligands has yet been identified for  
 CC the proteins of the invention. The orphan GPCRs may be used in the  
 CC identification of their endogenous ligands, and to screen potential GPCR  
 CC agonists and antagonists for use as pharmaceutical agents. The proteins  
 CC may also be used in the study of GPCR-mediated signalling cascades, and  
 CC to elucidate their precise role in normal and diseased human conditions.  
 CC Nucleic acid encoding human orphan GPCRs may be used for tissue  
 CC localisation analysis to provide information about their

CC function in healthy and pathological states.

XX Sequence 415 AA:

Query Match 99.1%; Score 1252; DB 21; Length 415;  
Best Local Similarity 99.6%; Pred. No. 4.5e-141;  
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKLNASWYIOOKLEDPQKHNLSTEEYLAFLGPRRSHFPLPVSVYPIFVGVIGN 60  
DB 4 MEKLNASWYIOOKLEDPQKHNLSTEEYLAFLGPRRSHFPLPVSVYPIFVGVIGN 63  
QY 61 VLVCVLILOHQAARKPTNTNYLFLSLAVSDLLVLLGMPLEYEMRWYPLFGVGCYFKT 120  
DB 64 VLVCVLILOHQAARKPTNTNYLFLSLAVSDLLVLLGMPLEYEMRWYPLFGVGCYFKT 123  
QY 121 ALFEYVCFASIISTITVSERYVAIIHPRAKLOSTRRALRLIGIVGFSVLSLPNTS 180  
DB 124 ALFEYVCFASIISTITVSERYVAIIHPRAKLOSTRRALRLIGIVGFSVLSLPNTS 183  
QY 181 IHGIRFHPNGSLVPGSATCTVIKPMIYNFTIOVTSFLFYLLPMTVISLVYLMALRY 240  
DB 184 IHGIRFHPNGSLVPGSATCTVIKPMIYNFTIOVTSFLFYLLPMTVISLVYLMALRL 243

RESULT 8  
AAB02830  
ID AAB02830 standard; Protein; 415 AA.

AC AAB02830;  
DT 22-AUG-2000 (first entry)

DE Human G protein coupled receptor hRue protein SEQ ID NO:12.

KW Human; G protein coupled receptor; GPCR; transmembrane receptor;  
KW identification; agonist; screening; therapeutic; pharmaceutical;  
KW mutant.

XX Homo sapiens.

PN WO200022131-A2.

XX PD 20-APR-2000.

XX PF 13-OCT-1999; 99WO-US24065.

XX PR 13-OCT-1998; 98US-0170496.

XX PR 12-NOV-1998; 98US-0108029.

XX PR 20-NOV-1998; 98US-0109213.

XX PR 27-NOV-1998; 98US-0110060.

XX PR 16-FEB-1999; 99US-0120416.

XX PR 26-FEB-1999; 99US-0121852.

XX PR 12-MAR-1999; 99US-0123944.

XX PR 12-MAR-1999; 99US-0123945.

XX PR 12-MAR-1999; 99US-0123946.

XX PR 12-MAR-1999; 99US-0123948.

XX PR 12-MAR-1999; 99US-0123949.

XX PR 12-MAR-1999; 99US-0123951.

XX PR 28-MAY-1999; 99US-0136436.

XX PR 28-MAY-1999; 99US-0136437.

XX PR 28-MAY-1999; 99US-0136439.

XX PR 28-MAY-1999; 99US-0137127.

XX PR 28-MAY-1999; 99US-0137131.

XX PR 30-JUN-1999; 99US-0137567.

XX PR 27-AUG-1999; 99US-0141448.

XX PR 03-SEP-1999; 99US-0151114.

XX PR 29-SEP-1999; 99US-0152524.

XX PR 29-SEP-1999; 99US-0156633.

XX PR 29-SEP-1999; 99US-0156555.

XX PR 29-SEP-1999; 99US-0156634.

XX PA (AREN-) ARENA PHARM INC.

XX Behan DP, Lehmann-Brunsmma K, Chalmers DT, Chen R, Dang HT;  
PI Core M, Liaw CW, Lin I, Lowitz K, White C;  
XX WPI: 2000-317986/27.  
DR N-PSDB: AAA46022.

PT Non-endogenous, human G protein-coupled receptors for screening  
PT receptor, inverse or partial agonists useful as therapeutic agents -  
XX Example 1; Page 86-88; 187pp: English.

CC The present invention describes transmembrane receptors, preferably  
CC human G protein coupled receptors (GPCR), for which the endogenous  
CC ligand is unknown (orphan GPCR receptors). More specifically the present  
CC invention relates to non-endogenous, constitutively activated versions  
CC of a human GPCR. These non-endogenous human GPCRs can be useful for  
CC the direct identification of candidate compounds as receptors agonists,  
CC inverse agonists or partial agonists for use as pharmaceutical agents.  
CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in  
CC the exemplification of the present invention.

XX Sequence 415 AA:

Query Match 99.1%; Score 1252; DB 21; Length 415;  
Best Local Similarity 99.6%; Pred. No. 4.5e-141;  
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKLNASWYIOOKLEDPQKHNLSTEEYLAFLGPRRSHFPLPVSVYPIFVGVIGN 60

DB 4 MEKLNASWYIOOKLEDPQKHNLSTEEYLAFLGPRRSHFPLPVSVYPIFVGVIGN 63

QY 61 VLVCVLILOHQAARKPTNTNYLFLSLAVSDLLVLLGMPLEYEMRWYPLFGVGCYFKT 120

DB 64 VLVCVLILOHQAARKPTNTNYLFLSLAVSDLLVLLGMPLEYEMRWYPLFGVGCYFKT 123

QY 121 ALFEYVCFASIISTITVSERYVAIIHPRAKLOSTRRALRLIGIVGFSVLSLPNTS 180

DB 124 ALFEYVCFASIISTITVSERYVAIIHPRAKLOSTRRALRLIGIVGFSVLSLPNTS 183

QY 181 IHGIRFHPNGSLVPGSATCTVIKPMIYNFTIOVTSFLFYLLPMTVISLVYLMALRY 240

DB 184 IHGIRFHPNGSLVPGSATCTVIKPMIYNFTIOVTSFLFYLLPMTVISLVYLMALRL 243

RESULT 9  
AA52992  
ID AA52992 standard; Protein; 415 AA.

AC AA52992;

DT 21-FEB-2000 (first entry)

DE Human neurotensin-like receptor.

KW Human; neurotensin-like receptor; NLR; G-protein coupled receptor;  
KW central nervous system; anesthesia; analgesia.

XX Homo sapiens.

PN WO9955732-A1.

XX PD 04-NOV-1999.

XX PF 15-APR-1999; 99WO-SE00598.

XX PR 24-APR-1998; 98SE-0001455.

XX PA (ASTR-) ASTRA PHARMA INC.

XX PA (ASTR-) ASTRA AB.

XX Ahmad S, Cao J, O'Donnell D, Walker P;

DR WPI; 2000-052803/04.  
 DR N-PSDB; AA233297.  
 XX  
 PT Novel neurotensin-like receptor. useful for identifying agents for  
 PT producing anaesthesia or analgesia -  
 XX  
 PS Claim 1; Fig 2; 46pp; English.  
 XX  
 CC The present sequence represents a human G-protein coupled receptor  
 CC neurotensin-like receptor (NLR). The NLR polynucleotide and protein  
 CC can be used to isolate compounds that bind, (ant)agonise or alter  
 CC the activity or expression of the NLR. The NLR is a G-protein coupled  
 CC receptor which is expressed in the central nervous system and shares  
 CC homology with human neurotensin receptor. The receptors can be used in  
 assays to identify agents for producing anaesthesia and analgesia.  
 QY Sequence 415 AA:  
 Query Match 99.1%; Score 1252; DB 21; Length 415;  
 Best Local Similarity 99.6%; Pred. No. 4.5e-141;  
 Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MEKIQNASWITQQKLEDPFGKHLSNTEEYLAFLCGPRRSHFFLPVSVVYPIFVGVIGN 60  
 DB 4 MEKIQNASWITQQKLEDPFGKHLSNTEEYLAFLCGPRRSHFFLPVSVVYPIFVGVIGN 63  
 QY 61 VLVCVLVILQHQAKTPTNYLFLSLAVSDLLVLLGMPLEVEYEMMRNPFLFGPGCYEKT 120  
 DB 64 VLVCVLVILQHQAKTPTNYLFLSLAVSDLLVLLGMPLEVEYEMMRNPFLFGPGCYEKT 123  
 QY 121 ALPEYVCFASILITTVSVRRYVAIILHPRAKLOSTRRALRILGIVGFSVLSLPTS 180  
 DB 124 ALPEYVCFASILITTVSVRRYVAIILHPRAKLOSTRRALRILGIVGFSVLSLPTS 183  
 QY 181 IHGIKFHYFPNGSLVPSGATCTVIKPMIYNFIIOVTSFLFYLLPMTVISLVLYLALRV 240  
 DB 184 IHGIKFHYFPNGSLVPSGATCTVIKPMIYNFIIOVTSFLFYLLPMTVISLVLYLALRV 243  
 RESULT 10  
 AAG63353  
 ID AAG63353 standard; Protein; 415 AA.  
 XX  
 AAG63353;  
 DT 15-OCT-2001 (first entry)  
 XX  
 DE Amino acid sequence of a human TGR-1 protein.  
 XX  
 KW TGR-1; neuromedin U; hypertension; stress disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157524-A1.  
 PD 09-AUG-2001.  
 XX  
 PF 02-FEB-2001; 2001WO-JP00746.  
 PR 04-FEB-2000; 2000JP-0032773.  
 PR 24-FEB-2000; 2000JP-0052252.  
 PR 30-MAR-2000; 2000JP-0097896.  
 PR 19-JUN-2000; 2000JP-0187536.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;  
 PI Okubo S;  
 XX  
 DR WPI; 2001-488917/53.  
 DR N-PSDB; AAH43072.  
 XX  
 PT Identifying predicted or actual structures of two or more members of a

PT chemical or physical library by mass spectrometry comprising  
 PT correlating molecular mass measurements of two or more members with a  
 PT shared chemical history -  
 XX  
 PS Claim 1; Page 77-79; 95pp; Japanese.  
 XX  
 CC The present sequence represents a human TGR-1 protein. The specification  
 CC describes a method of screening a compound, which is capable of binding  
 CC properties of neuromedin U to TGR-1. The method is useful for screening  
 CC preventatives and remedies for hypertension, stress diseases, etc..  
 CC TGR-1 antagonists are also useful for treating the same diseases.  
 XX  
 QY Sequence 415 AA:  
 Query Match 99.1%; Score 1252; DB 22; Length 415;  
 Best Local Similarity 99.6%; Pred. No. 4.5e-141;  
 Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MEKIQNASWITQQKLEDPFGKHLSNTEEYLAFLCGPRRSHFFLPVSVVYPIFVGVIGN 60  
 DB 4 MEKIQNASWITQQKLEDPFGKHLSNTEEYLAFLCGPRRSHFFLPVSVVYPIFVGVIGN 63  
 QY 61 VLVCVLVILQHQAKTPTNYLFLSLAVSDLLVLLGMPLEVEYEMMRNPFLFGPGCYEKT 120  
 DB 64 VLVCVLVILQHQAKTPTNYLFLSLAVSDLLVLLGMPLEVEYEMMRNPFLFGPGCYEKT 123  
 QY 121 ALPEYVCFASILITTVSVRRYVAIILHPRAKLOSTRRALRILGIVGFSVLSLPTS 180  
 DB 124 ALPEYVCFASILITTVSVRRYVAIILHPRAKLOSTRRALRILGIVGFSVLSLPTS 183  
 QY 181 IHGIKFHYFPNGSLVPSGATCTVIKPMIYNFIIOVTSFLFYLLPMTVISLVLYLALRV 240  
 DB 184 IHGIKFHYFPNGSLVPSGATCTVIKPMIYNFIIOVTSFLFYLLPMTVISLVLYLALRV 243  
 RESULT 11  
 AAG63366  
 ID AAG63366 standard; Protein; 415 AA.  
 XX  
 AAG63366;  
 DT 15-OCT-2001 (first entry)  
 XX  
 DE Amino acid sequence of a human TGR-1 protein.  
 XX  
 KW TGR-1; neuromedin U; hypertension; stress disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157524-A1.  
 PD 09-AUG-2001.  
 XX  
 PF 02-FEB-2001; 2001WO-JP00746.  
 PR 04-FEB-2000; 2000JP-0032773.  
 PR 24-FEB-2000; 2000JP-0052252.  
 PR 30-MAR-2000; 2000JP-0097896.  
 PR 19-JUN-2000; 2000JP-0187536.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;  
 PI Okubo S;  
 XX  
 DR WPI; 2001-488917/53.  
 DR N-PSDB; AAH43075.  
 XX  
 PT Identifying predicted or actual structures of two or more members of a  
 PT chemical or physical library by mass spectrometry comprising  
 PT correlating molecular mass measurements of two or more members with a  
 PT shared chemical history -  
 XX

PS Disclosure; Page 88; 95pp; Japanese.

CC The present sequence represents a human TGR-1 protein. The specification describes a method of screening a compound, which is capable of binding CC properties of neuromedin U to TGR-1. The method is useful for screening CC preventatives and remedies for hypertension, stress diseases, etc.. CC TGR-1 antagonists are also useful for treating the same diseases.

XX Sequence 415 AA;

Query Match 99.1%; Score 1252; DB 22; Length 415;  
Best Local Similarity 99.6%; Pred. No. 4.5e-141;  
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFLOQASMTIYQOKLDDPPOKHLNSTEYLAFLCGRRSHFFLPVSVVYPIPVGVIGN 60  
DB 4 MEFLOQASMTIYQOKLDDPPOKHLNSTEYLAFLCGRRSHFFLPVSVVYPIPVGVIGN 63  
QY 61 VLVGLVILQHOAKRTPNTNYLFLSLAVSDLLVLLGMPLEVEYEMRMNYPFLFGVGCYFRT 120  
DB 64 VLVGLVILQHOAKRTPNTNYLFLSLAVSDLLVLLGMPLEVEYEMRMNYPFLFGVGCYFRT 123  
QY 121 ALPEWCFASILSTTVSVERVYAILHPPRAKQSTRRALRLGIWGFVSFLSPNTS 180  
DB 124 ALPEWCFASILSTTVSVERVYAILHPPRAKQSTRRALRLGIWGFVSFLSPNTS 183  
QY 181 IHGIRKHPNPGSLVPGSACIVIKPMWYINFTIQVTSFLFYLLPMTVISLVLYMALRV 240  
DB 184 IHGIRKHPNPGSLVPGSACIVIKPMWYINFTIQVTSFLFYLLPMTVISLVLYMALRL 243

RESULT 12  
AAG64297  
ID AAG64297 standard; Protein; 415 AA.

XX AAG64297;

XX 21-SEP-2001 (first entry)

XX Human GTP-binding protein-coupled receptor GPRV39.

XX GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;  
XX muscular; urinary; circulatory; anorectic; human; guanosine triphosphate;  
XX G-protein.

XX Homo sapiens.

XX WO200148189-A1.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-JP09409.

XX 28-DEC-1999; 99JP-0375152.

XX 31-MAR-2000; 2000JP-0101339.

XX 23-MAY-2000; 2000JP-0155978.

XX (HELI-) HELIX RES INST.

XX Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;

XX Sugiyama T;

XX WPI; 2001-425663/45.

XX N-PSDB; AAH49526.

XX Family of guanosine triphosphate binding protein coupled receptors and

XX genes encoding them for treatment and prevention of diseases associated

XX with these receptors

XX Claim 1; Pages 100-103; 137pp; Japanese.

XX The present sequence is the protein sequence for a human guanosine

XX triphosphate (GTP)-binding protein-coupled receptor. The receptor is

CC useful for the investigation, diagnosis, treatment and prevention of  
CC diseases associated with GTP-binding protein-coupled receptors, including  
CC neurological, circulatory, digestive system, immune system, muscle and  
CC urinary system disorders. GTP-binding proteins are also known as  
CC G-proteins.

XX Sequence 415 AA;

Query Match 99.1%; Score 1252; DB 22; Length 415;  
Best Local Similarity 99.6%; Pred. No. 4.5e-141;  
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEFLOQASMTIYQOKLDDPPOKHLNSTEYLAFLCGRRSHFFLPVSVVYPIPVGVIGN 60  
DB 4 MEFLOQASMTIYQOKLDDPPOKHLNSTEYLAFLCGRRSHFFLPVSVVYPIPVGVIGN 63  
QY 61 VLVGLVILQHOAKRTPNTNYLFLSLAVSDLLVLLGMPLEVEYEMRMNYPFLFGVGCYFRT 120  
DB 64 VLVGLVILQHOAKRTPNTNYLFLSLAVSDLLVLLGMPLEVEYEMRMNYPFLFGVGCYFRT 123  
QY 121 ALPEWCFASILSTTVSVERVYAILHPPRAKQSTRRALRLGIWGFVSFLSPNTS 180  
DB 124 ALPEWCFASILSTTVSVERVYAILHPPRAKQSTRRALRLGIWGFVSFLSPNTS 183  
QY 181 IHGIRKHPNPGSLVPGSACIVIKPMWYINFTIQVTSFLFYLLPMTVISLVLYMALRV 240  
DB 184 IHGIRKHPNPGSLVPGSACIVIKPMWYINFTIQVTSFLFYLLPMTVISLVLYMALRL 243

RESULT 13  
AAE03629  
ID AAE03629 standard; Protein; 415 AA.

XX AAE03629;

XX 07-AUG-2001 (first entry)

XX Human G-protein coupled receptor, SNORF72.

XX Human; G-protein coupled receptor; SNORF72; neuromedin U neuropeptide;  
XX NMU; inflammation; arthritis; autoimmune disease; septicemia; psychotic;  
XX mental retardation; transplant rejection; neurological disorder; anxiety;  
XX respiratory disorder; depression; schizophrenia; dementia; obesity; pain;  
XX gastrointestinal disorder; hypertension; epilepsy; diabetes;  
XX ischaemia; stroke; cancer; sexual disorder; circadian disorder; anorexia;  
XX dermatological; psoriasis; Parkinson's disease; nausea; bulimia; allergy;  
XX Alzheimer's disease; AIDS; hormonal disorder; memory disorder; migraine;  
XX cardiovascular disorder; renal disorder; bone disease; delirium; asthma;  
XX Cushing's disease; dysmenorrhea; antiangiinal; cytostatic; osteoporosis;  
XX metabolic disorder; behavioural disorder; Addison's disease; dyskinesia;  
XX tranquiliser; antiulcer; antiaddictive.

XX Homo sapiens.

XX Location/Qualifiers

XX Key 44..71

XX Domain /Label= Transmembrane\_domain

XX Domain /Label= Transmembrane\_domain

XX Domain /Label= Transmembrane\_domain

XX Domain /Label= Transmembrane\_domain

XX Domain /Label= Transmembrane\_domain

XX Domain /Label= Transmembrane\_domain

XX Domain /Label= Transmembrane\_domain

XX Domain /Label= Transmembrane\_domain

XX WO200144297-A1.

XX 21-JUN-2001.

XX 13-DEC-2000; 2000MO-US3787.  
 XX PF  
 XX 17-DEC-1999; 99US-0466435.  
 PR 25-APR-2000; 2000US-0558099.  
 PR 30-JUN-2000; 2000US-0609146.  
 XX  
 PA (SYNA-) SYNAPTIC PHARM CORP.  
 XX  
 PI Bonini JA, Lerman GS, Quan Y, Ogozalek K;  
 DR WPI; 2001-380240/41.  
 DR N-PSDB; AAD08008.  
 XX  
 A purified mammalian SNORF62 or SNORF72 receptor protein for  
 identification of compounds to treat e.g. inflammation, arthritis,  
 autoimmune diseases, transplant rejection, AIDS, cancer -  
 Claim 12; Fig 4; 256pp; English.

XX The invention relates to human G-protein coupled receptors, SNORF62 and  
 CC SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72  
 CC receptors are specific for neuropeptide U (NMU) neuropeptides, hence they  
 CC are also known as NMU receptors. The agonist and antagonist of NMU  
 CC receptors are useful for treating an abnormality in a subject that is  
 CC alleviated by decreasing or increasing the activity of NMU receptor.  
 CC The NMU receptors serve as a valuable tool for designing drugs which are  
 CC useful for treating various pathophysiological conditions such as  
 CC inflammation, arthritis, autoimmune diseases, transplant rejection,  
 CC graft vs host disease, bacterial, fungal, protozoan and viral infections,  
 CC septicemia, AIDS, pain, psychotic and neurological disorders, including  
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory  
 CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,  
 CC eating/body weight disorders including obesity, bulimia, diabetes,  
 CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular  
 CC disorders, ischemia, stroke, cancers, sexual disorders, circadian  
 CC disorders, renal disorders, bone diseases including osteoporosis, benign  
 CC prostatic hypertrophy, gastrointestinal disorders, nasal congestion,  
 CC dermatological disorders such as psoriasis, allergies, Parkinson's  
 CC disease, Alzheimer's disease, acute heart failure, angina disorders,  
 CC delirium and dyskinesias such as Huntington's disease. They can also be  
 CC used to regulate steroid hormone disorders, epinephrine release  
 CC disorders, electrolyte balance disorders, endocrine disorders, memory  
 CC disorders, somatosensory disorders, metabolic disorders, behavioural  
 CC disorders, drug addiction, migraine, Addison's disease, Cushing's  
 CC disease, prevent miscarriage, induce labour or to treat dysmenorrhoea.  
 CC The present sequence is human G-protein coupled receptor, SNORF72.

XX Sequence 415 AA:  
 SQ

Query Match 99.1%; Score 1252; DB 22; Length 415;  
 Best Local Similarity 99.6%; Pred. No. 4.5e-141;  
 Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MERKQNASWITYOQKLEDDPFOKHLNSTEYLAFLCGPRRSHPFLPVSVVYPIFVGVIGN 60  
 DB 4 MERKQNASWITYOQKLEDDPFOKHLNSTEYLAFLCGPRRSHPFLPVSVVYPIFVGVIGN 63

QY 61 VLVCLVYLQQAQAKTPNTNYLFLSLAVSDLLVLLGMPLEVEYEMKRNYPFLFGPGCYEFT 120  
 DB 64 VLVCLVYLQQAQAKTPNTNYLFLSLAVSDLLVLLGMPLEVEYEMKRNYPFLFGPGCYEFT 123

QY 121 ALPETVCFASILTITTVSVRRYVALHPFRAKLOSTRRRRLRLITGTYWGSVLFSLPNTS 180  
 DB 124 ALPETVCFASILTITTVSVRRYVALHPFRAKLOSTRRRRLRLITGTYWGSVLFSLPNTS 183

QY 181 IHGIKFHYFPNGSLVPGSACIVIKPMWYINFTIQVTSFLFYLLPMTVIVSVLYLMA 240  
 DB 184 IHGIKFHYFPNGSLVPGSACIVIKPMWYINFTIQVTSFLFYLLPMTVIVSVLYLMA 243

RESULT 14  
 AAB67802

ID AAB67802 standard; Protein; 415 AA.  
 XX  
 AC AAB67802;  
 XX  
 DT 29-JUN-2001 (first entry)  
 XX  
 DE Amino acid sequence of G-protein coupled receptor IGS4A long version.  
 XX  
 KW Human: G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;  
 KW nervous system disorder; psychiatric disorder; Parkinson's disease;  
 KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;  
 KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;  
 KW motility disorder; myocardial infarction; hypertension; dyslipidemia;  
 KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;  
 KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;  
 KW gynecological disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200125269-A2.  
 XX  
 PD 12-APR-2001.  
 XX  
 PF 25-SEP-2000; 2000MO-EP09584.  
 XX  
 PR 24-SEP-1999; 99EP-0203140.  
 PR 24-SEP-1999; 99NT-1013140.  
 PR 28-JUL-2000; 2000EP-0202683.  
 PR 31-JUL-2000; 2000US-0222047.  
 XX  
 PA (SOLV ) SOLVAY PHARM BV.  
 XX  
 PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;  
 DR WPI; 2001-273568/28.  
 DR N-PSDB; AAF0322.  
 XX  
 PT New G-protein coupled receptors and the polynucleotides encoding them,  
 PT useful for preventing, ameliorating or correcting nervous system  
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain  
 PT or cancers -  
 XX  
 XX Claim 18; Page 77-79; 102pp; English.

XX The present sequence represents the long version of a human G-protein  
 CC coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,  
 CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful  
 CC for preventing, ameliorating or correcting dysfunctions or diseases.  
 CC These diseases include peripheral nervous system, psychiatric and central  
 CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal  
 CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or  
 CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,  
 CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,  
 CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility  
 CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,  
 CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,  
 CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are  
 CC effective with regard to disorders of the nervous system, including the  
 CC central and peripheral nervous systems, disorders of the gastrointestinal  
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or  
 CC gastrointestinal system, or immunological disease. The IGS4 polynucleotides  
 CC are useful as diagnostic reagents for detecting under-expression,  
 CC overexpression or altered expression of IGS4.  
 XX  
 SQ Sequence 415 AA:  
 QY

Query Match 99.1%; Score 1252; DB 22; Length 415;  
 Best Local Similarity 99.6%; Pred. No. 4.5e-141;  
 Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MERKQNASWITYOQKLEDDPFOKHLNSTEYLAFLCGPRRSHPFLPVSVVYPIFVGVIGN 60  
 DB 4 MERKQNASWITYOQKLEDDPFOKHLNSTEYLAFLCGPRRSHPFLPVSVVYPIFVGVIGN 63

```

OY 61 VLVCVILIOHQAAMKPTNYLFLSLAVSDLLVLLGMPLEVEEMRNYPFLFGPVGCYFRT 120
DB 64 VLVCVILIOHQAAMKPTNYLFLSLAVSDLLVLLGMPLEVEEMRNYPFLFGPVGCYFRT 123
OY 121 ALFETVCFASIISTITVSERYVAIILHPRAKIQSTRRALRLIGIWMGSVLSLPNTS 180
DB 124 ALFETVCFASIISTITVSERYVAIILHPRAKIQSTRRALRLIGIWMGSVLSLPNTS 183
OY 181 IHGIRKHYPPNGSLVPGSATCTYIKPMWYNFIQVTSFLFYLLPMTVLSVLYLMALRV 240
DB 184 IHGIRKHYPPNGSLVPGSATCTYIKPMWYNFIQVTSFLFYLLPMTVLSVLYLMALRV 243

RESULT 15
AAB67804
ID AAB67804 standard: Protein: 415 AA.
AC AAB67804;
XX
XX
XX 29-JUN-2001 (first entry)
DE
XX Amino acid sequence of a G-protein coupled receptor IG54B long version.
XX
XX Human; G-protein coupled receptor; IG54; IG54A; IG54B; schizophrenia;
XX nervous system disorder; psychiatric disorder; Parkinson's disease;
XX episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
XX cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
XX motility disorder; myocardial infarction; hypertension; dyslipidemia;
XX gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
XX inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
XX gynecological disorder.
XX
XX Homo sapiens.
XX
XX WO200125269-A2.
XX
XX 12-APR-2001.
XX
XX 25-SEP-2000; 2000MO-EP09584.
XX
XX 24-SEP-1999; 99EP-0203140.
XX 24-SEP-1999; 99NL-1013140.
XX 28-JUL-2000; 2000EP-0202683.
XX 31-JUL-2000; 2000US-0222047.
XX
XX (SOLV ) SOLVAY PHARM BV.
XX
XX Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX
XX WPI: 2001-273568/28.
XX N-PSDB; AAF80324.
XX
XX New G-protein coupled receptors and the polynucleotides encoding them,
XX useful for preventing, ameliorating or correcting nervous system
XX disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
XX or cancers
XX
XX Claim 19; Page 85-86; 102pp; English.
XX
XX
XX The present sequence represents the long version of a human G-protein
XX coupled receptor designated IG54B. IG54 exists in two polymorphic forms,
XX IG54A and IG54B. The IG54 receptors and IG54 polynucleotides are useful
XX for preventing, ameliorating or correcting dysfunctions or diseases.
XX These diseases include peripheral nervous system, psychiatric and central
XX nervous system disorders (e.g. schizophrenia, episodic paroxysmal
XX anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
XX stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
XX myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
XX gastrointestinal disorders (e.g. inflammatory bowel disease or motility
XX disorders), osteoporosis, inflammations, infections (e.g. bacterial,
XX fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
XX sepsis or gynecological disorders. Agonists or antagonists of IG54 are

```

```

CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IG54 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IG54.
XX
XX Sequence 415 AA:
SQ
Query Match 99.1%; Score 1252; DB 22; Length 415;
Best Local Similarity 99.6%; Pred. No. 4,5e-141;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 MERLONASWITVOOKLEDDPQKHUNSTEEYLAFLCGRSHFPLPVSVVYPIFFVVCIGN 60
DB 4 MERLONASWITVOOKLEDDPQKHUNSTEEYLAFLCGRSHFPLPVSVVYPIFFVVCIGN 63
OY 61 VLVCVILIOHQAAMKPTNYLFLSLAVSDLLVLLGMPLEVEEMRNYPFLFGPVGCYFRT 120
DB 64 VLVCVILIOHQAAMKPTNYLFLSLAVSDLLVLLGMPLEVEEMRNYPFLFGPVGCYFRT 123
OY 121 ALFETVCFASIISTITVSERYVAIILHPRAKIQSTRRALRLIGIWMGSVLSLPNTS 180
DB 124 ALFETVCFASIISTITVSERYVAIILHPRAKIQSTRRALRLIGIWMGSVLSLPNTS 183
OY 181 IHGIRKHYPPNGSLVPGSATCTYIKPMWYNFIQVTSFLFYLLPMTVLSVLYLMALRV 240
DB 184 IHGIRKHYPPNGSLVPGSATCTYIKPMWYNFIQVTSFLFYLLPMTVLSVLYLMALRV 243

```

Search completed: January 17, 2003, 05:04:03  
Job time : 83 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 04:30:56 : Search time 41 Seconds  
(without alignments)  
567.428 Million cell updates/sec

Title: US-09-684-725-2

Perfect score: 1263

Sequence: 1 MEKLNASWVYQKLEDPQ.....LLPMTVISLVYLMALRVSI 242

Working table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir73:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	338	26.8	424	2 JH0164	neurotensin recept
2	332.5	26.3	378	2 T13816	hypothetical prote
3	325.5	25.8	418	2 S29506	neurotensin recept
4	303.5	24.0	418	2 A88013	protein K10B4.4 [1
5	281	22.2	416	2 S68822	neurotensin recept
6	281	22.2	564	2 A38271	serotonin receptor
7	265	21.0	393	2 A39251	thyrotropin receptor
8	259	20.5	398	2 JN0708	thyrotropin-releas
9	258	20.4	411	2 I56444	thyrotropin-relea
10	258	20.4	412	2 S23436	thyrotropin recep
11	255.5	20.4	402	2 I56595	neurokinin 2 recep
12	255.5	20.2	392	2 S65693	opioid receptor mu
13	255	20.2	398	2 A57510	opioid receptor mu
14	253.5	20.1	400	2 I56553	mu opiate receptor
15	252.5	20.0	369	2 B41795	somatostatin recep
16	248	19.6	369	2 JC2083	somatostatin recep
17	247.5	19.6	369	2 A45291	somatostatin recep
18	243.5	19.3	384	2 S20303	neurokinin 1 recep
19	243	19.2	363	2 I57940	somatostatin recep
20	242.5	19.2	385	2 S55524	neurokinin 3 recep
21	242.5	19.2	398	2 I56517	mu opiate receptor
22	241.5	19.1	384	1 S00516	neurokinin 2 recep
23	241.5	19.1	359	2 S15403	angiotensin II rec
24	240.5	19.0	442	1 DYL4D2	dopamine receptor
25	240	19.0	359	2 JH0621	angiotensin II rec
26	240	19.0	359	2 JC2134	angiotensin II rec
27	239.5	18.9	345	2 T24659	hypothetical prote
28	239	18.9	359	2 JC1104	angiotensin II rec
29	239	18.9	367	2 I49022	kappa opiate recep

30	239	18.9	367	2 JC2421	opioid receptor ho
31	239	18.9	367	2 I56520	G protein-coupled
32	238.5	18.9	346	2 S29248	somatostatin recep
33	238.5	18.9	369	2 D41795	somatostatin recep
34	238.5	18.9	370	2 S43087	orphan opioid rece
35	238.5	18.9	452	2 A34916	neurokinin 3 recep
36	238	18.8	359	2 S44425	angiotensin II rec
37	238	18.8	359	2 A42656	angiotensin II rec
38	237.5	18.8	398	1 J01059	neurokinin 2 recep
39	237.5	18.8	511	2 C56849	dopamine receptor
40	237	18.8	444	1 DYM5D2	dopamine receptor
41	237	18.8	444	1 S08146	dopamine receptor
42	237	18.8	460	2 A32605	serotonin receptor
43	236.5	18.7	387	2 JC5949	galanin receptor 2
44	235	18.6	359	2 A48857	angiotensin II rec
45	234	18.5	359	2 JC1194	angiotensin II rec

## ALIGNMENTS

```

RESULT 1
JH0164
neurotensin receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Mar-2000
C:Accession: JH0164
R:Tanaka, K.; Masu, M.; Nakanishi, S.
A:Title: Structure and functional expression of the cloned rat neurotensin receptor.
A:Reference number: JH0164; PMID:90297956; PMID:1694443
A:Accession: JH0164
A:Molecule type: mRNA
A:Residues: 1-424 <TRAN>
C:Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor.
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:65-87/Domain: transmembrane #status predicted <TM1>
F:97-121/Domain: transmembrane #status predicted <TM2>
F:144-165/Domain: transmembrane #status predicted <TM3>
F:189-210/Domain: transmembrane #status predicted <TM4>
F:236-260/Domain: transmembrane #status predicted <TM5>
F:309-330/Domain: transmembrane #status predicted <TM6>
F:348-372/Domain: transmembrane #status predicted <TM7>
F:4,38,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          26.8%; Score 338; DB 2; Length 424;
Best Local Similarity 35.1%; Pred. No. 1,3e+22;
Matches 72; Conservative 50; Mismatches 71; Indels 12; Gaps 5;

QY 45 VSVVYVPIFYVGVIGNVCLVILQH--QAMKPTNYVFLASVLDLVLLGLMPLEVY 101
      | : : : | | | | : : : : : : : : : : | | | | | | | | : : |
Db 67 VTAIYILFLVYGVGVNSVTAFTLARKKSLQSLSTVYHNGSLALSDLLILLAMPVELY 126
      | : : : | | | | : : : : : : : : : : | | | | | | | | : : |
QY 102 E-WWRNYPFLFGPVGC--YFKTALPFTVCFASTLSTTVSVRYVYVLLHPPRAKLOSTR 157
      | : : : | | | | : : : : : : : : : : | | | | | | | | : : |
Db 127 NFIWVHHPWAFGAGCGYTF--LRDACTYATLAWNASVEERYLATICHPEFAKTLMRSR 183
      | : : : | | | | : : : : : : : : : : | | | | | | | | : : |
QY 158 RRALRLIGIVWGFVSFLPNTSLHGKIFHYFPNGSLVPGSACCTVTKPMIYNFIIOVT 217
      | : : : | | | | : : : : : : : : : : | | | | | | | | : : |
Db 184 SRRKRFSAIWLASALLATIMLFMTGLQNR--SGDGTNPGGLVCTPIVDPAATVAKVIVQVN 241
      | : : : | | | | : : : : : : : : : : | | | | | | | | : : |
QY 218 SFLFLPLPMTVISLVYLMALRVSI 242
      | : : : | | | | : : : : : : : : : : | | | | | | | | : : |
Db 242 TFWSEFLPMLVISILNTVIANKLTV 266
      | : : : | | | | : : : : : : : : : : | | | | | | | | : : |

RESULT 2
T13816
hypothetical protein C48C5.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000

```

C:Accession: T15816  
 R:Ravello, A.  
 Submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of C. elegans cosmid C48C5.  
 A:Reference number: Z18410  
 A:Accession: T15816  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-378 <FAV>  
 A:Cross-references: EMBL:U09994; NID:g1055102; PID:g1055105; PIDN:AAB37017.1; GSPDB:GN00  
 A:Experimental source: strain Bristol N2; clone C48C5  
 C:Gene: CESP:C48C5.1  
 A:Map position: X  
 A:Insertions: 53/3; 87/2; 136/2; 169/3; 209/3; 231/3; 259/3; 286/1; 327/1  
 C:Superfamily: adenosine receptor A1

Query Match 26.3%; Score 332.5; DB 2; Length 378;  
 Best Local Similarity 33.6%; Pred. No. 3.6e-22;  
 Matches 87; Conservative 37; Mismatches 98; Indels 37; Gaps 6;

QY 20 QKHLNTEER---LALICGPRRH-----FLLPVSVVYPIF 53  
 DB 3 QACLNTEEDCDDCLAFNCPIYSHSESEKACYEHCIFSKRALDQVTLKVTALYIFIF 62  
 DB 54 VGVIGAVLVCLVTLIOHQAOKTPTNYLFLSLAVSDLLVLLGMPLEYEMRNYPFLFGR 113  
 DB 63 LVGIGVITTCVLLVKKRPMKTHASMTLMNLAVSDVTLVGLPEEYMMNMNOIPWPFDP 122  
 QY 114 VGVYFATLFEVYCFASILSTTVYVERVYVALIHP-FRAKLOSTRRALRILGIWGFVS 172  
 DB 123 YICMKALIAETTSVSLITLILFAIERVYAVACHPLFMKVOPKRNIGITGTFWFSI 182  
 QY 173 LFSPLNTHIGIKF--HYFP---NGSLVPGSACTVT-----KPMIYNTIIOVTSFLFY 222  
 DB 183 LCAPPRFIIHRADIMKSWGTDRIIPVSKSKMCMIAMVEPKLASTFKILFHSAIAAF 242  
 QY 223 LRPVTIVSLVLYMALRV 241  
 DB 243 ALPLFTIVILYARACKVS 261

RESULT 3  
 S29506  
 neurotensin receptor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 17-Mar-2000  
 C:Accession: S29506  
 R:Vilta, N.; Laurent, P.; Lefort, S.; Chalton, P.; Dumont, X.; Kaghad, M.; Gully, D.; Le F  
 FEBIS Lett. 317, 139-142, 1993  
 A:Title: Cloning and expression of a complementary DNA encoding a high affinity human ne  
 A:Reference number: S29506; MUID:93154505; PMID:8381365  
 A:Accession: S29506  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-418 <VIT>  
 A:Cross-references: EMBL:X70070; NID:g35020; PIDN:CAA49675.1; PID:g35021  
 C:Superfamily: vertebrate rhodopsin  
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 25.8%; Score 325.5; DB 2; Length 418;  
 Best Local Similarity 32.6%; Pred. No. 1.7e-21;  
 Matches 86; Conservative 49; Mismatches 76; Indels 53; Gaps 10;

QY 17 DDPQKHLNTEEL-----AFICGPRRS-----HFLPVSVVYPIFV 55  
 DB 17 DDPQKHLNTEEL-----AFICGPRRS-----HFLPVSVVYPIFV 76  
 QY 56 GVIGAVLVCLVTLIOHQAOKTPTNYLFLSLAVSDLLVLLGMPLEYE-MMRNYPFLF 111  
 DB 77 GYIGAVLVCLVTLIOHQAOKTPTNYLFLSLAVSDLLVLLGMPLEYE-MMRNYPFLF 136  
 QY 112 GPGVC---YFKTALFETVCFASILSTTVYVERVYVALIHPFRAKLOSTRRALRILGIW 168

DB 137 GDAGCGRGYF---LRDCTATATANASLSVERLALCHPFAKTLMSRSRTKFTSAIW 193  
 QY 169 GFSYLSLP-----NTSIGHIKFHYFNGSLV--PGSATCTVIKPMIYNTIIOVTS 218  
 DB 194 LASALLIVPLFLFMGEQNRSDGQH-----AGGLVCTPIHTATV-----KVIQVNT 241  
 QY 219 FLEYLPMYIVSVLYYIMALRVSI 242  
 DB 242 FMSFIEPMYIVSVLYYIMALKTV 265

RESULT 4  
 A88013  
 protein K10B4.4 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: A88013  
 R:anonymous; The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;  
 A:Accession: A88013  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-418 <STO>  
 A:Cross-references: GB:chr.II; PIDN:AB71009.1; PID:g429475; GSPDB:GN00020; CESP:K10B  
 A:Note: similar to family I of G-protein coupled receptors  
 C:Gene: K10B4.4  
 A:Map position: 2

Query Match 24.0%; Score 303.5; DB 2; Length 418;  
 Best Local Similarity 30.2%; Pred. No. 1.6e-19;  
 Matches 76; Conservative 54; Mismatches 89; Indels 33; Gaps 6;

QY 22 HUNSTEELAFLOGR--KSHFLLPVSVVYPIFVGVIGAVLVCLVTLIOHQAOKTPTNY 79  
 DB 9 NSEITEYVSLTIGERCQASGAGIPIVITIGTFLGLGFCNICTCIVIANKNSMHNPTNY 68  
 QY 80 YLFSLAVSDLLVLLGMPLEYE-MMRNYPFLFPGVYCFKATLFEVYCFASILSTTVTS 138  
 DB 69 YLFSLAVSDLLVLLGMPLEYE-MMRNYPFLFPGVYCFKATLFEVYCFASILSTTVTS 128  
 QY 139 VERYVALIHPFRAKLOSTRRALRILGIWGFVSLEPNTSIHGT-----184  
 DB 129 FERWLAICHPLRSKIRSTLMRAWVLIITLAWTISFVCALEPIAFVIOINKLPEDAKYQW 188  
 QY 185 --KEHFEP---NGSLVPGSAT-----CTVIKPMIYNTIIOVTSF-LFYLIPMTV 228  
 DB 189 TNKSFFAVGVINNRIFPVSTDGIIVLHTEFCAMNOSRPDOOKMIIIFAFVFEVIPAIA 248  
 QY 229 ISVLYYIMALRV 240  
 DB 249 IYIMTAHIAVOL 260

RESULT 5  
 S68822  
 neurotensin receptor 2, levocabastine-sensitive - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 04-Dec-1997 #sequence\_revision 12-Dec-1997 #text\_change 20-Jun-2000  
 C:Accession: S68822  
 R:Chalton, P.; Vilta, N.; Kaghad, M.; Guillemot, M.; Bonnin, J.; Delpesch, B.; Le Fur, G  
 FEBIS Lett. 386, 91-94, 1996  
 A:Title: Molecular cloning of a levocabastine-sensitive neurotensin binding site.  
 A:Reference number: S68822; MUID:96228041; PMID:8647296  
 A:Accession: S68822  
 A:Molecule type: mRNA  
 A:Residues: 1-416 <CHA>  
 A:Cross-references: GB:X07121; NID:g1483579; PIDN:CAA65787.1; PID:g1483580



A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-263 <HIN>  
A:Accession: S50152  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 267-398 <HI2>  
R:Duhrstie, S.M.; Taylor, P.L.; Anderson, L.; Cook, J.; Eidne, K.A.  
A:Title: Cloning and functional characterisation of the human TRH receptor  
A:Reference number: I38356; MUID:94063224; PMID:8243797  
A:Accession: I38356  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-398 <RES>  
A:Cross-references: EMBL:X72089; NID:g4440155; PIDN:CAA50979.1; PID:g4440156  
C:Genetics:  
A:Gene: GDB:TRMR  
A:Cross-references: GDB:228955; OMIM:188545  
A:Map position: 8q23-8q23  
C:Superfamily: adenosine receptor A1  
C:Keywords: G protein-coupled receptor; receptor; transmembrane protein  
F:29-51/Domain: Transmembrane #status predicted <TM1>  
F:62-83/Domain: Transmembrane #status predicted <TM2>  
F:101-121/Domain: Transmembrane #status predicted <TM3>  
F:146-168/Domain: Transmembrane #status predicted <TM4>  
F:194-215/Domain: Transmembrane #status predicted <TM5>  
F:267-288/Domain: Transmembrane #status predicted <TM6>  
F:297-319/Domain: Transmembrane #status predicted <TM7>

Query Match	20.5%	Score	259;	DB 2;	Length	398;			
Best Local Similarity	32.4%	Pred	No. 1.4e-15;						
Matches	66;	Conservative	47;	Mismatches	65;	Indels	26;	Gaps	7;

```

QY 45 VSVVYVPFLV-VGNIQGNVLVCVLILQHQAMKTPNNVYLFSLAIVSDLVLVL-LGMPL---E 99
      |::| |::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |::|::| |::|::|
Db 27 VTILLVLIICGLGIVGNIMVVLVMRTKTHMRTPNICYLVSIAVADIMVLYAAGLEPNITDS 86

```

**OY**    100 YEMERNRPFLGPGVCYFKTALFETVCFASILSTTTVSERYVALHPFAKLQSTRR    159  
      : |     :: |||      |    ||     ::||:|||     |    :  
**Dd**    87 IYGSW----VGXVGLCTIYLQLGINASSCSITAFIERIYAICHPIKAQFLCFSR    141

```

QY 160 ALRILGIWGFSLVLESLPNPMSIHGIRKHFEPNGSLVPGSATC-----TVIKPMWLYNFI 213
      | : | | : : : : : : : | : : : : |
Db 142 AKKIIFWVAFTSLYCMLEWFLDDLNISTYKDAIVI-----SCGYKISRNYYPILMDFC 197

```

QY	214	IQVTSFLFYLLPMIVISLVLYIMA	237
		::   : :    :	
Db	198	V-----FYVPMILATVILGYFIA	215

RESULT 9  
I56444  
cortrophin-releasing hormone receptor - mouse  
species: Mus sp. (mouse)

C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 11-Jan-2000  
 C:Accession: I56444  
 J:SELLAR, R.E.; TAYLOR, P.L.; LAMB, R.F.; ZABAVNIK, J.; ANDERSON, L.; EIDNE, K.A.  
 J: MOL. ENDOCRINOI., 10, 199-206, 1993  
 A:Title: Functional expression and molecular characterization of the thyrotrophin-releasing hormone receptor A1  
 A:Reference number: I56444; MUID:93249585; PMID:8387312  
 A:Accession: I56444  
 A:Status: preliminary; translated from GR/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-411 <RES>  
 A:Cross-references: GB:560053; NID:g300151; PIDN:AAB26491.1; PID:g300152  
 A:Superfamily: adenosine receptor A1

Query Match	20.4%	Score 258;	DB 2;	Length 411;
Best Similarity	30.4%;	Pred. No. 1.8e-15;		
Match Local 72;	Conservative 49;	Mismatches 80;	Indels 36;	Gaps 8;

QY 15 LEDPQKHLNSTEYLAFLCGPRSHFELPVSVVPIFV---GVIGNVLCLVILQH 70

```

Db      1  MENFVSELNÖTEL-----PQVAALEXYÖVITLLVIGGLGGINIMVLTVMRE 53
Oy      71  OAMKTPMYVYFSLVAASDLLVLL-LGML---EYEMKMYRPLBEPVCNGCYKRTALFETV 126
Db      54  KHMRAINCYLVSIALDMLVAAGLPNTDISYGSW-----VGVYSCLCITTYÖYUG 108
Oy      127  CFASLITTTVSVERYAAILHPEFAKQSTRRALRLIGVYFSLPLMPTSHIGIKF 186
Db      109  INASCSSTATTIRRYAICHPIKAQCLCFESAKKIIIFVNAFTSYOMLMEPLDLMI 168
Oy      187  HYPNGSLVPSASAC-----TVIKPMWIIINFLIÖVTSLEFYLLPMTVSVLYILMA 237
Db      169  STYDAIYI-----SCGYKISRNNYSPIYLMDFEV-----FYVMPNITAFVYIGFA 215

```

RESULT 10  
S23436

thyrotropin-releasing hormone receptor  
N:Alternate names: thyrotropin-releasing hormone receptor

C:Species: Rattus norvegicus (Norway rat)  
C:Date: 13-Jan-1995 #sequence revision 13

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 11-Jan-2000  
C:Accession: S23436; I53279; A49168; PQ0326

R; de La Peña, P.; Delgado, L.M.; del Camino, D.; Barros, F. Biochem. J. 284, 891-899, 1992

A1:Title: Cloning and expression of the thyrotropin-releasing hormone receptor from GH  
A1:Reference number: S23436; MUID:92322017; PMID:1377915

A/Accession: S23436  
A/Molecule type: mRNA

A;Residues: 1-412 <PEN>  
A;Cross-references: EMBL:X64630; NID:q57394; PIDN:CAA45913.1; PID:q57395

R; Kimura, N.; Arai, K.; Sahara, Y.; Suzuki, H.; Kimura, N. *Endocrinology* 134, 432-440, 1994

A:Title: Estradiol transcriptionally and posttranscriptionally up-regulates thyrotrop  
A:Reference number: 153279; PMID:9410223; PMID:8275956

A/Accession: I53279  
A/Status: preliminary: translated from GB/EMBL/DDBJ

A:Residues: 1-412 <RE

A: Cross-references: GB:D17469; NID:g464199; PIDN:BA04289.1; PID:g464200  
R: Zhao, D.; Yang, J.; Jones, K. F.; Gerald, C.; Suzuki, Y.; Nogata, D. C.

Attila: Molecular cloning of a complementary deoxyribonucleic acid encoding the Abv  
Endocrinology 130, 3529-3536, 1992

A: Accession: A49168  
A: Reference number: A49168; MUID: 92283212; PMID: 131778787

A:Accession: A45108  
A:Status: preliminary  
A:Molecule type: mpna

A;Residues: 1-12, 'D', 14-290, 'K', 292-412 <ZHA>

A: Note: sequence extracted from NCBI backbone (NCBIN:104788, NCBIIP:104795)

Riyamada, M.; Monden, T.; Satoh, T.; Izuka, M.; Murakami, M.; Iriuchijima, T.; Mori, Biochem. Biophys. Res. Commun., 184, 367-372, 1992

A:Title: Differential regulation of thyrotropin-releasing hormone receptor mRNA level  
A:Reference number: P00326; MUID:92231953; PMID:1373613

A:Accession: PQ0326  
A:Molecule type: mRNA

A;Residues: 30-58, 'P', 60-222, 'T', 224-261 <YAM>  
A;Experimental source: strain wister

A; Note: the authors translated the codon ACA for residue 88 as Ala  
C; Superfamily: adenosine receptor A1

C; Keywords: G protein-coupled receptor; transmembrane protein F; 1-22/Domain: transmembrane #status predicted <TM1>

F;32-54/Domain:	transmembrane	#status	predicted	<TM2>
F;71-92/Domain:	transmembrane	#status	predicted	<TM3>

F;116-140/Domain:	transmembrane	#status	predicted	<TM
F:165-186/Domain: <td>transmembrane</td> <td>#status</td> <td>predicted</td> <td>&lt;TM</td>	transmembrane	#status	predicted	<TM

Query Match 20.4% Score 258: DB 2: Length 412:

Best Local Similarity	30.48	Pred. No. 1.8e-15;
Matches 72: Conservative	49	Mismatches 80
		Indels 36
		Gaps 9

	12; conserved	45; mismatches	60; indels	36; gaps
15 LEDPDKH NSTFEYAF CGPDSHSEFL DVSUVPBTEW- ---GVICNNVQLVITQH 700				

10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

```

Db      1 MENEIVSELNQL-----PPQVAVALLEQVVTIILLVVICGLIGINIMVLVMKRT   53
Oy      71 QAMKTPTNYLFLSLAVSDLVL--LGMPL--EYEMRMNPFLGPGVCYEKTALEFETV   126
       1 : 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db      54 KMRRTATNCYLTVSLAVDLMLVLAAGLEPRLPNTIDTSIGSW----YGVGCCLITYLYLG   108
Oy     127 CPASLTSTTVEVERVYVILHPRAKLQSTRRRALRIIGVWGFSVLFSLPNTSIHGKE   186
       1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db     109 IWASSCSITAFITEKRIAICHPIKAOFCTESRRAKIIFWAPFTSYCMAMPFLDINT   168
Oy     187 HYFPGSLPGSANTC-----TVIKPMWNIIOVTSFEFLPLMFIYSLVLYLMA   237
       1 : 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db     169 STYKDAIVI---SCGYKRISRNRYSPILMDFGV-----FYVMPLMILAIVLYGRFA   215
SUBT    11
156595
neurokinin 2 receptor - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 20-Apr-2000
C:Accession: I56595
R:Baronny, D., Little, J., Thomas, C.; Powell, S.; Downey-Jones, M.; Graham, A.
A:Recept. Res. 14, 359-421, 1994
A>Title: Isolation and characterization of neurokinin A receptor cDNAs from guinea-Pig
A:Reference number: I56595; MUID:95182423; PMID:7877137
A:Accession: I56595
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-402 <RES>
A:Cross-references: GB:S76253; NID:g913274; PIDN:AAB33553.1; PID:g913275
C:Superfamily: neurokinin 1 receptor

```

RESULT 12  
S65693  
Opioid receptor mu variant MOR1A - human  
C.Species: Homo sapiens (man)  
C.Date: 12-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 05-Nov-1999  
C.Accession: S65693; S51216  
R.Bare, L.A.; Mansson, E.; Yang, D.  
Submitted to the EMBL Data Library, July 1994  
A.Description: Expression of two variants of the human mu opioid receptor mRNA in SK-N-SH cells  
A.Reference number: S65693  
A.Accession: S65693  
A.Molecule type: mRNA  
A.Residues: 1-392<BAR>  
A.Cross-references: EMBL:U12569; NID:9607911; PIDN:AAB60354.1; PID:9607912  
R.Bare, L.A.; Mansson, E.; Yang, D.  
FEBS Lett. 354, 213-216, 1994  
A.Title: Expression of two variants of the human mu opioid receptor mRNA in SK-N-SH cells  
A.Reference number: S51215; MUID:95046336; PMID:7957926  
A.Accession: S51216  
A.Molecule type: mRNA

[illegible]

RESULT 13

A57510  
mu opioid receptor - mouse

C:Species: Mus musculus (house mouse)  
C:Date: 08-Dec-1995 #sequence\_revision 08-Dec-1995 #text\_change 24-Nov-1999  
C:Accession: A57510; I48665; S66513; I49300  
R:Kaufman, D.L.; Keith Jr., D.E.; Anton, B.; Tian, J.; Megendzo, K.; Newman, D.; Tran  
J. BIOL. Chem. 270, 15877-15883, 1995  
A:Title: Characterization of the murine mu opioid receptor gene.  
A:Reference number: A57510; MUID:95318184; PMID:7797593  
A:Accession: A57510  
A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA  
A:Residues: 1-398 <KAT>

A:Cross-references: GB:U19380  
R:Min, B.H.; Augustin, L.B.; Felsheim, R.F.; Fuchs, J.A.; Loh, H.H.  
Proc. Natl. Acad. Sci. U.S.A. 91, 9081-9085, 1994  
A:Title: Genomic structure analysis of promoter sequence of a mouse mu opioid receptor  
A:Reference number: I48665; MUID:94377496; PMID:8090773  
A:Accession: I48665  
A:Status: translated from GR/EMBL/DBDT

A:Molecule type: DNA  
A:Residues: 1-398 <RES>

A:Cross-references: EMBL:U10561; NID:9555696; PIDN:AA860673.1; PID:9565069  
R:Rossi, G.C.; Pan, Y.X.; Brown, G.P.; Pasternak, G.W.  
FEBS Lett. 369, 192-196, 1995  
A:Title: Antisense mapping the MOR-1 opioid receptor: evidence for alternative splicing  
A:Reference number: I49300; MUID:95377399; PMID:7649256  
A:Accession: S66513  
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA  
A:Residues: 1-398 <ROS>

A:Cross-references: EMBL:U26915; NID:91055230; PIDN:AA81170.1; PID:91055231  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995  
C:Genetics:  
A:Gene: MOR-1  
A:Introns: 95/2; 213/1; 386/3  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembr

Query Match 20.2%; Score 255; DB 2; Length 398;  
Best Local Similarity 30.7%; Pred. NO. 3.le-15;  
Matches 70; Conservative 42; Mismatches 84; Indels 32; Gaps 8;

34 CGPRR-----SHFFLP-----VSVVYVIFVGVIGNVLCVLIIHQAMKTP 76



Fri Jan 17 09:12:12 2003

us-09-684-725-2.rpr

Page 7

QY 203 VIKP---MMIVNIIQVTSSEFLYLLPMETVISLVLYLMALRV 240  
: | | | : : : : : : : :  
Db 195 INMPGESGAWYTGRIF-YFILGLFVLPLTITICTLCTLFILIKV 235

Search completed: January 17, 2003, 05:07:32  
Job time : 47 secs

THIS PAGE BLANK (L-3PT0)





FT	CAROHYD	38	38	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CAROHYD	42	42	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	DISULFID	142	225	BY SIMILARITY.
FT	LIPID	368	368	PALMITATE (POTENTIAL).
SO	SEQUENCE	424 AA:	47054 MW;	A9C2FEAF8D9BCD3 CRC64;
Query Match				
Best Local Similarity		26.8%;	Score 338;	DB 1;
Matches		72;	Conservative	50;
			Mismatches	71;
			Indels	12;
			Gaps	
QY	45	VSVVYVPIRVGVIGNVLCVLILOH---	QAMKPTNYLFLANSDDLVLGMPLEVY	101
DB	67	VTAIVLALVFLVGVGSMVTAFLARKKSLOSQSTVHYHGLSTALSDLLILLAMPVELY	126	
QY	102	E--MNRNYFLGVFGVC--YFKTALFEYGCPSILISITTVSERYVAIIHPRAKIQSR	157	
DB	127	NETIWNHHPRAFDAGCRGYF---LRDACTYATFALWASLSERYATLCHPRAKIAMS	183	
QY	158	RRLALRLIGVWGPVSFLSLPNTSHIGIKFHPNPSLVGSACTGVYIKPMWYNFLIQYT	217	
DB	184	STTKRKISAIWLASALLAPMLFETMGLQNR--SGDGTHEGGLVCPYIDTATVAVKVIQVN	241	
QY	218	SFLFVLLPMTVTSVLVYLMALRFSI	242	
DB	242	TFMSFLPMLVTISLTVYIAKLTV	266	
RESULT 2				
MTLR_HUMAN				
ID	MTLR_HUMAN	STANDARD:	PRT:	412 AA.
AC	043193:			
DT	15-DEC-1998	(Rel. 37, Created)		
DT	15-DEC-1998	(Rel. 37, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Motilin receptor (G protein-coupled receptor GPR38).			
OS	MTLR1 OR MTLR OR GPR38.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euteria; Primates; Carnivora; Hominiina; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM A):			
RX	MEDLINE=98110578; PubMed=9441746;			
RX	McGee K.K., Tan C.P., Palyha O.C., Liu J., Feighner S.D.,			
RA	Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.,			
RT	"Cloning and characterization of two human G protein-coupled receptor			
RT	genes (GPR38 and GPR39) related to the growth hormone secretagogue			
RT	and neurotensin receptors.";			
RL	Genomics 46:426-434(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORMS A AND B).			
RX	MEDLINE=99316084; PubMed=10381885;			
RX	Feighner S.D., Tan C.P., McGee K.K., Palyha O.C., Hreniuk D.L.,			
RA	Pong S.-S., Austin C.P., Figueroa D., MacNeil D., Cascieri M.A.,			
RT	Nargund R., Bakshi R., Abramovitz M., Stocco R., Kargman S.,			
RT	O'Neill G., van Der Ploeg L.H.T., Evans J., Patchett A.A., Smith R.G.,			
RA	Howard A.D.;			
RT	"Receptor for motilin identified in the human gastrointestinal			
RT	system.";			
RT	Science 284:2184-2188(1999).			
RN	[3]			
RN	SEQUENCE FROM N.A. (ISOFORM A).			
RA	Wall M.;			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	FUNCTION.			
RX	MEDLINE=21219832; PubMed=11322507;			
RA	Smith R.G., Leonard R., Bailey A.R.T., Palyha O.C., Feighner S.D.,			
RA	Tan C.P., McGee K.K., Pong S.-S., Griffin P.R., Howard A.D.;			
RT	"Growth hormone secretagogue receptor family members and ligands.";			
RL	Endocrine 14:9-14(2001).			
CC	-1- FUNCTION: Receptor for motilin.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			

CC	- I - ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are produced by alternative splicing.
CC	- I - TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE MARROW.
CC	- I - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/or-send-an-email-to-license@sib-stb.ch">http://www.isb-sib.ch/announce/or-send-an-email-to-license@sib-stb.ch</a> ).
CC	or send an email to license@sib-stb.ch)
CC	-----
DR	EMBL: AF034632; AAC26081.1; -
DR	EMBL: AL137000; CAC19107.1; -
DR	Genew; HGNC:4495; GPCR38.
DR	MIM; 602885; -
DR	InterPro; IPR000276; GPCR_Rhodpsn.
DR	Pfam; PF00001; 7tm1.1;
DR	PRINTS; PR00237; GPCRRHODPSN.
DR	PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
DR	PROSITE; PS00262; G-PROTEIN_RECP_F1_2; 1.
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;
KW	Alternative splicing.
FT	DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 56 56 1 (POTENTIAL).
FT	DOMAIN 37 74 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 75 94 2 (POTENTIAL).
FT	DOMAIN 95 112 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 113 134 3 (POTENTIAL).
FT	DOMAIN 135 157 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 158 178 4 (POTENTIAL).
FT	DOMAIN 179 246 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 247 270 5 (POTENTIAL).
FT	DOMAIN 271 298 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 299 320 6 (POTENTIAL).
FT	DOMAIN 321 334 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 335 358 7 (POTENTIAL).
FT	DOMAIN 359 412 CYTOPLASMIC (POTENTIAL).
FT	DISELFID 111 235 BY SIMILARITY.
FT	CARBOHYD 6 6 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD 192 192 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	VANSPPLIC 301 412 LUNVARLICWLPVGRIRIINTEDSRMVFSGYENIVAL
SO	SEQUENCE 412 AA; 45344 MW; CIJFP616501ZDER3 CRC64;
Query Match	26.4%; Score 333; DB 1; Length 412;
Best Local Similarity	31.8%; Pred. No. 2,7e-17;
Matches 78; Conservative 51; Mismatches 70; Indels 46; Gaps 5;	
OY	CGPRSHFFELVSVVYVIFEVGVGNVLVCCLYLIIQHAMKTPFNYYFLSLAVSDLVLL 93
DB	30 CSPFPLGALVPATACDLCELYVGSGNVAYMLIGRYRDMFTTNVLGSMAVSDLLI-L 88
OY	LGMPELEVEMRNPFELRGPGCYEFTKLFEPTVCFASILTSTTVSVERYVALILHPRAKL 153
DB	89 LGLPEDLVLRMKRSRMWVGPLLRLSLTLVYGSGCYATILLHTALTALSVERIYLAICRPRLARV 148
OY	154 QSTRRALRIIGIYWGSVLESLPNTSIHGIKEHYFPNGSLVPG-----SATCTVIK 205
DB	149 LVTRRRRVALLAVLMAVALLSAGPFLIVGE--QDPGISVIVPGTANTARIASSPLASSP 206
OY	PMWI-----YNFIQVTSFIEFLLMPIYISV 231
DB	207 PLWLNRAPPSPSGPETAFAALFSSRECRPSPAQDGLAKVLMWVTVAIFYF-LDFLCIST 265
OY	232 LYLTLM 236

Db 266 LYGLI 270

## RESULT 3

NTSL\_MOUSE  
ID NTSL\_MOUSE STANDARD: PRT: 424 AA.  
AC 088319;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Neurotensin receptor type 1 (NT-R-1).  
GN NTSL OR NTSL.  
OS Mus musculus (Mouse).  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_Taxid=10090;  
[1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Snider J., Sano H., Ohta M.;  
RT Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.  
RL -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. HIGHEST TO TACHYKININ RECEPTORS.  
-----  
CC This SWISS-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: AB017027; BAA33013.1; -  
DR MGD: MG1:97386; Nstr.  
DR InterPro: IPR000276; GPCR\_Rhodopsn.  
DR InterPro: IPR003985; NT1\_rec.  
DR InterPro: IPR003984; NT\_rec.  
DR Pfam: PFO001; 7tm.1; 1.  
DR PRINTS: PRO0237; GPCR\_Rhodopsn.  
DR PRINTS: PRO1479; NEUROTENSINR.  
DR PRINTS: PRO1480; NEUROTENSINR.  
DR PROSITE: PS00237; G\_PROTEIN\_RECP\_FL\_1; 1.  
DR PROSITE: PS0262; G\_PROTEIN\_RECP\_FL\_2; 1.  
KW G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Phosphorylation; Lipoprotein; Palmitate.  
FT DOMAIN 1 64  
FT TRANSMEM 65 87  
FT DOMAIN 88 96  
FT TRANSMEM 97 121  
FT DOMAIN 122 143  
FT TRANSMEM 144 165  
FT DOMAIN 166 188  
FT TRANSMEM 189 210  
FT DOMAIN 211 235  
FT TRANSMEM 236 260  
FT DOMAIN 261 308  
FT TRANSMEM 309 330  
FT DOMAIN 331 348  
FT TRANSMEM 349 372  
FT DOMAIN 373 424  
FT CARBOHYD 4 4  
FT CARBOHYD 38 38  
FT CARBOHYD 42 42  
FT CARBOHYD 211 211  
FT DISULFID 141 224  
FT LIPID 388  
SQ SEQUENCE 424 AA: 47216 MW: 8E9A723171A48711 CRC64;

Query Match 26.3%; Score 332.5; DB 1; Length 424;

Best Local Similarity 32.1%; Pred. No. 3e-17;

Matches 80; Conservative 55; Mismatches 85; Indels 29; Gaps 7;

OY 4 LONASNIYOOKLEDPPOKHLNSTEFLALCGPRSHFLEPSVYVPIFVGVGNVLV 63  
DB 36 LNSGNSSESEILEPNSNDVNTDIVSKVL-----VYAVLALFVVGTVGNSVT 84  
OY 64 CLVILH---QAMKPTNYFLSLAVSDLLVLLGPLEVE-MMRNPFLEPVVC--- 116  
DB 85 AFLTARRKSLQSLQSTVHYHLGSLADLLILLAMPVELNFIVWHHPAFEDACRG 144  
OY 117 YEKTALEFYCFASILSTTVSEYVVAIIHPERAKLQSTRRLRLILGIVGFSYLSF 176  
DB 145 YF---LRDACTVATLNVASLSVERLALCHPKATKMSRSTKFKISAWLASALLAV 201  
OY 177 PNTSHIGKFEHPNGSL---VPGSATCYIKPMIYNFIQVTSFLFLLPMTVTSVL 233  
DB 202 PMLFTMGLO-----NRSADQHPGCHVCTPTVDATVVKVIVQVTFMPSFLPMLITSLN 256  
OY 234 YLMALRVS 242  
DB 257 TVIANKLTV 265

## RESULT 4

NTSL\_HUMAN  
ID NTSL\_HUMAN STANDARD: PRT: 418 AA.  
AC P30989; Q9H4T5; Q9H4H1;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Neurotensin receptor type 1 (NT-R-1) (High-affinity leucobastine-insensitive neurotensin receptor) (NTRH).  
GN NTSL OR NTRR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NCBI\_Taxid=9606;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-93154505; PubMed-8381365;  
RA Vita N., Laurent P., Lefort S., Chalou P., Dumont X., Kaghad M., Gully D., Le Fur G., Ferrara P., Caput D.;  
RT "Cloning and expression of a complementary DNA encoding a high affinity human neurotensin receptor."  
RT FEBS Lett. 317:139-142(1993).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-21638749; PubMed-11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babage A.K., Baguley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blake S.E., Bridgman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R., Coulson A., Coville G.J., Dearden R., Dhani P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Levasailho M.H., Leverisa M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A., Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillips B.J.C.T., Prachalam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showman R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E., Swan R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,

RA Wilmington J., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.:  
RT "The DNA sequence and comparative analysis of human chromosome 20.",  
RL Nature 414:865-871(2001)  
CC  
CC -I- FUNCTION: RECEPTOR FOR THE TRIDCAPEPTIDE NEUROTENSIN. IT IS  
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-  
CC CALCIUM SECOND MESSENGER SYSTEM.  
CC  
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.  
CC  
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC HIGHEST TO TACHYKININS RECEPTORS.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC	-
DR EMBL; X70070;	CAA49675.1;-
DR EMBL; AL357033;	CAC14923.1;-
DR EMBL; AL035669;	CAGC12747.1;-
DR PIR; S29506;	S29506.
HSSP; P02689;	1F88..
DR Genew; HGNC; R039;	MNSR.L
MM; M62651;	--
R InterPro; IPR000276;	GPCR_Rhodopsn.
N InterPro; IPR003985;	NTI_rec.
DR InterPro; IPR003984;	NT_rec.
Pfam; PF00001;	7tm_1-2
DR Pfam; PR01479;	NEUROTENSINR.
DR PRINTS; PR01479;	NEUTOTENSN.R
DR PROSITE; PS00237;	G_PROTEIN_REC.F1.1; 1.
DR PROSITE; PS00262;	G_PROTEIN_REC.F1.2; 1.
KW G-protein coupled receptor:	Transmembrane; Glycoprotein,
KM Phosphorylation; Lipoprotein;	Palmitate.
FT DOMAIN	1 63 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	64 86 CYTOSOLASMIC (POTENTIAL).
FT DOMAIN	87 95 2 (POTENTIAL).
FT TRANSMEM	96 120 EXTRACELLULAR (POTENTIAL).
FT DOMAIN	121 142 3 (POTENTIAL).
FT TRANSMEM	143 164 CYTOSOLASMIC (POTENTIAL).
FT DOMAIN	165 187 4 (POTENTIAL).
FT TRANSMEM	188 209 EXTRACELLULAR (POTENTIAL).
FT DOMAIN	210 234 5 (POTENTIAL).
FT TRANSMEM	235 259 CYTOSOLASMIC (POTENTIAL).
FT DOMAIN	260 303 6 (POTENTIAL).
FT TRANSMEM	304 325 EXTRACELLULAR (POTENTIAL).
FT DOMAIN	326 363 CYTOSOLASMIC (POTENTIAL).
FT TRANSMEM	344 367 7 (POTENTIAL).
DOMAIN	368 418 CYTOPLASMIC (POTENTIAL).
F CATABOHD	4 4 N-LINKED (GLCNAC. .)
FT CARBOHYD	37 37 N-LINKED (GLCNAC. .)
FT CARBOHYD	41 41 POT-N LINKED (GLCNAC. .)
LIPID DISULFD	141 224 BY SIMILARITY.
LIPID	383 383 PALMITATE (POTENTIAL).
CONFLECT	200 200 T -> A (LN REF. 2).
SEQUENCE	418 AA: 46288 MW: BBBDIECR2BEBE390 CRC64;

Query Match	25.8%;	Score 325.5;	DB 1;	Length 418;	.
Best Local Similarity	32.6%;	Pred. No. 9.3e-17;			.
Matches 86;	Conservative 49;	Mismatches 76;	Indels 53;	Gaps 10	

[illegible]

Db 137 GDACGRGYF--LRDACTAATATNVAASLVERLALCHPFRAKTLMRSRTRKKTSAIM 193

QY 169 GFSVLESLP-----NTSHGKIKPIFYFNGSYV--PSGATCTYKPKMIIYNFIIQYTS 218

Db 194 LASALLIYPMLEFTEGONRSDGSH-----AGGLYCIPTIHTATV-----KVIQVONT 241

QY 219 FLFLILEPRTVLSVLYLMLALRST 242

Db 242 FMSFLFPPWVIVSLNTLIANKLV 265

RESULT 5	NTNR2_RAT	NTNR2_RAT	STANDARD;	PRT;	416 AA.
ID	NTNR2_RAT				
AC	Q63384;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Neurotensin receptor type 2 (NT-R-2) (High-affinity levocastaine-sensitive neurotensin receptor).				
GN	NTSR2 OR NTNR2				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eultheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Hypothalamus;				
RC	MEDLINE=66228041; Pubmed=8647296;				
RA	Chalon P., Vita N., Kaqhad M., Guillemont M., Bonin J.,				
RA	Delpech B., Le Fur G., Ferrara P., Caput D.;				
RT	"Molecular cloning of a levocastaine-sensitive neurotensin binding site.";				
RL	FEBS Lett. 386:91-94(1996).				

- 1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
- 1- SUBCELLULAR LOCATION: Integral membrane protein.
- 1- TISSUE SPECIFICITY: ABUNDANT IN CORTIX AND HYPOTHALAMUS, AND LOWER LEVELS SEEN IN THE HEART AND INTESTINE.
- 1- DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY IN 7-DAY-OLD BRAIN AND EXPRESSION DECREASES PROGRESSIVELY UNTIL ADULTHOOD (35-DAY-OLD BRAIN).
- 1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. HIGHEST TO TACHYKININS RECEPTORS.

This SWISS-PROT entry is copyrighted and is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (<http://www.isb-sib.ch/announce/>). For send an email to [license@sib-ch.ch](mailto:license@sib-ch.ch).

DR	InterPro: IPR000276; GPCR_Rhodopsn.	
DR	InterPro: IPR003986; NF2_Rec.	
DR	InterPro: IPR003984; NF_Rec.	
DR	Pfam: PF00001; 7tm1; 1.	
DR	PRINTS: PR01479; NEUROKENSINR.	
DR	PRINTS: PR01481; NEUROKENSZR.	
DR	PROSITE: PS00237; G_PROTEIN_RECPE_F1_1; 1.	
DR	PROSITE: PS00262; G_PROTEIN_RECPE_F1_2; 1.	
KW	G-protein coupled receptor; Transmembrane; Lipoprotein; Extracellular (POTENTIAL).	Palmitate
FT	DOMAIN 1	32
FT	TRANSMEM 33	55
FT	DOMAIN 56	64
FT	TRANSMEM 65	87
FT	DOMAIN 88	109
FT	TRANSMEM 110	131
FT	DOMAIN 132	154
FT	TRANSMEM 155	176
FT	DOMAIN 177	216

RA K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
Dudin A.E., Garg N.S., Gelbart W.M., Grasser K.,  
Foster C., Gabriellian A.E.,

Matches	79:	Conservative	36:	Mismatches	73:	Indels	28:	Gaps	9:
---------	-----	--------------	-----	------------	-----	--------	-----	------	----

```

OY 27 EBYLA--FLGCRSHFPLVSVVYPIFVVGIVGNVLVCLIIHQAMKPTNYLFSL 84
DB 148 EEFAGEFVPLPLTSIF---VSIVLLIYILGVGNVNLVCLAVCNVRLKRPNCNLLVSL 204
OY 85 AVSDLLVLLGMPL-----EYEMRNRYPLFGPGCYKRTALFETVCASILSTTVSVE 140
DB 205 ALSDLCVALLVPMPLLVLEKRN-----FCPLLCDIWSFEDVLCVASILNCAISVD 259
OY 141 RVVALIHFRKALOSTRRRLRIIGIYVGFVLSLPTMSIHGIKHFHPGSLVPGSAT 200
DB 260 RLATITKLEGVCKRTPRRMKCVGIWLAACISLPLLLTGNLHEDERGQPI----- 313
OY 201 CTVIKPMWIVNFIIQVTSFL--FILLPMTVLSVLY 234
DB 314 CTVCO-----NFAYOIVATLSFY-IPLSVLMFYVY 343

RESULT 7
GHSR_HUMAN STANDARD: PRT; 366 AA.
AC 092847; 092848; 096R37;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing peptide receptor) (GHRP) (Ghrelin receptor).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
RC TISSUE=Pituitary;
RX MEDLINE=96337998; PubMed=8688086;
RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P., Liberatore P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L., Palyna O.C., Anderson J., Patess P.S., Diaz C., Chou M., Liu K.K., Makke K.K., Pong S.-S., Chaung L.-Y., Elbrecht A., Daskevich M., Heavens R., Ridsby M., Stinathsinghji D.J.S., Dean D.C., Melillo D.G., Raebelt A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K., Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;
RA "A receptor in pituitary and hypothalamus that functions in growth hormone release.";
RT Science 273:974-977(1996).
RL [2]
RN SEQUENCE FROM N.A. (ISOFORMS 1A AND 1B).
RP MEDLINE=21255649; PubMed=11356716;
RA Petersen S., Rasch A.C., Peshorn M., Bell F.U., Schulte H.M.;
RT "Genomic structure and transcriptional regulation of the human growth hormone secretagogue receptor.";
RT Endocrinology 142:2649-2659(2001).
RN [3]
RN FUNCTION.
RN MEDLINE=2129832; PubMed=11322507;
RA Smith R.G., Leonard R., Bailey A.R.T., Palyna O.C., Feighner S.D., Tan C.P., McKee K.K., Pong S.-S., Griffin P.R., Howard A.D.;
RT "Growth hormone secretagogue receptor family members and ligands.";
RT Endocrine 14:9-14(2001).
RN [4]
RN FUNCTION.
RN MEDLINE=20067959; PubMed=10604470;
RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;
RT "Ghrelin is a growth-hormone-releasing acylated peptide from stomach.";
RT Nature 402:656-660(1999).
CC -i- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins. Stimulates growth hormone secretion. Binds also other growth hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6) as well as non-peptide, low molecular weight secretagogues (e.g. l-692,429, MK-0677, adenosine).
CC -i- SUBCELLULAR LOCATION: Integral membrane protein.
CC -i- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A (SHOWN HERE) AND 1B; ARE

```

```

CC CC PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND
CC CC SECRETAGOGUES.
CC CC -i- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
CC CC -i- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL: U60179; AAC50653.1; -
DR EMBL: U60181; AAC50654.1; -
DR EMBL: AF369786; AAK71539.1; -
DR EMBL: AF369786; AAK71540.1; -
DR GeneW: HGNC:4267; GHSR.
DR MIM: 601898; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Alternative splicing.
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 41 66 1 (POTENTIAL).
FT DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 73 96 2 (POTENTIAL).
FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 163 183 4 (POTENTIAL).
FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 235 5 (POTENTIAL).
FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 264 285 6 (POTENTIAL).
FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 303 326 7 (POTENTIAL).
FT DOMAIN 327 366 CYTOPLASMIC (POTENTIAL).
FT DISULFD 116 198 BY SIMILARITY.
FT CARBOHYD 13 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 27 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 266 289 AVVPARTICULPHVGRYIFSKS -> GSGORALRSLAG
FT FT PILSLCLPSL (IN ISOFORM 1B).
FT FT MISSING (IN ISOFORM 1B).
SQ SEQUENCE 366 AA; 41328 MM; DIB62710DA9DC0C6 CRC64;

Query Match 21.9%; Score 277; DB 1; Length 366;
Best Local Similarity 31.7%; Pred. No. 2.5e-13;
Matches 66; Conservative 47; Mismatches 69; Indels 26; Gaps 5;

OY 45 VSVYVPIFVVGIVGNVLVCLIIHQAMKPTNYLFSLAVSDLLVLLGMPLEVEWM 104
DB 46 VYATCVLAFFVGIAGNLLTLMVSRFRLTNYLISMAFSDLLFLIC-MPLDIVRLM 104
OY 105 RNYRPLFGPGCYKRTALFETVCASILSTTVSVERVALHFRKALOSTRRRLRIIL 164
DB 105 QREWNFGDLLCKLFQFVSECTYATVLTALSERFAICPLRAKVVTGKGRVRLVI 164
OY 165 GIWGFVLSLPTMSIHGIKHFHPGNS-----LVGSAFCVYIKPMWLYNF 212
DB 165 FVYMAVACASGPIFLVAVGEH-----NCTDPMWINECRPIEFVAVRSGLLTV--VW--- 216
OY 213 ITQVTSFLFYLLPMTVLSVLYMALRV 240
DB 217 -----SSIFFLPVFCILVLYSLIGRKL 239

RESULT 8
GHSR_RAT STANDARD: PRT; 364 AA.

```

AC 008725;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing peptide receptor) (GHRP) (Ghrelin receptor).  
 GN GHSR.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NX NCBI\_TaxId=10116;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 TISSUE-Pituitary;  
 MEDLINE=97246555; PubMed=9092793;  
 McKee K.K., Palya O.C., Feighner S.D., Hreniuk D.L., Tan C.P., Phillips M.S., Smith R.G., der Ploeg L.H.T., Howard A.D.;  
 "Molecular analysis of rat pituitary and hypothalamic growth hormone secretagogue receptors.";  
 Mol. Endocrinol. 11:415-423(1997).  
 RL 121  
 RP SEQUENCE OF 1-240 FROM N.A.  
 RC STRAIN-Mistar; TISSUE-Pituitary;  
 RX MEDLINE=98100386; PubMed=9437732;  
 RX Yokote R., Sato M., Matsubara S., Ohye H., Niimi M., Murao K., Takahara J.;  
 "Molecular cloning and gene expression of growth hormone-releasing peptide receptor in rat tissues.";  
 Peptides 19:15-20(1998).  
 RL 131  
 RP FUNCTION.  
 RX MEDLINE=20067959; PubMed=10604470;  
 RA Kojima M., Hosoda H., Date Y., Nakazato M., Matsuo H., Kangawa K.;  
 "Ghrelin is a growth-hormone-releasing acylated peptide from stomach.";  
 Nature 402:656-660(1999).  
 RL Nature 402:656-660(1999).  
 CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins. Stimulates growth hormone secretion. Binds also other growth hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6) as well as non-peptide, low molecular weight secretagogues (e.g. L-692,429, MK-0677, adenosine) (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U94321; AAC53156.1;  
 DR EMBL: AB001982; BAA21777.1; ALT\_INIT.  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm.1.1.  
 DR PRINTS: PR00237; GPCRHOPOPSN.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECPT\_FL\_1; 1.  
 DR PROSITE: PS50262; G\_PROTEIN\_RECPT\_FL\_2; 1.  
 KW 6-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 40  
 FT TRANSMEM 41 66  
 FT DOMAIN 67 72  
 FT TRANSMEM 73 96  
 FT DOMAIN 97 117  
 FT TRANSMEM 118 139  
 FT DOMAIN 140 162  
 FT TRANSMEM 163 183  
 FT DOMAIN 184 211  
 FT TRANSMEM 212 235  
 FT DOMAIN 236 262  
 FT TRANSMEM 264 285  
 FT DOMAIN 302  
 EXTRACTED FROM (POTENTIAL).  
 EXTRACTED FROM (POTENTIAL).

FT TRANSMEM 303 326 7 (POTENTIAL).  
 FT DOMAIN 327 364 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 115 197 BY SIMILARITY.  
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 364 AA: 40963 MW: DCEFS59BE061EEB9 CRC64:  
 Query Match 21.7% Score 274; DB 1; Length 364;  
 Best Local Similarity 31.7%; Pred. No. 4, 2e-13;  
 Matches 66; Conservative 47; Mismatches 69; Indels 26; Gaps 5;  
 QY 45 VSVYVPIEVGVIGNVLCVILQHQAMKTPNNYLSLAVSDLLVLLGMPLEYEM 104  
 DB 45 VTAICVAFVGVSGNLTMLVVSRRRLRTTNILKLSNAFSDLLIFLC-MPLDLVRLW 103  
 QY 105 RNPPLFPGVPCYFKTALFEFTVCASILSTTVSVVERVAIILHPRAKQSTRRLRL 164  
 DB 104 QYRPNWFGDLCKLFQFVSSECTYAVLTITLVSVERFAICPLRAKVVTGRVKLVI 163  
 QY 165 GIYMGFSVLFSLPNTSHGKIFHPNGS-----LVPGSATCVIKPMIYNF 212  
 DB 164 LVIMAVAFCSAGPIFVLGVENE---NGTDPRDTNECRATERAVRSGLLTVM--VWV--- 215  
 QY 213 IIQVTSFLPYLPMYIVSYLYMALRV 240  
 DB 216 -----SSVFFFLVFCITVYLSLGRKL 238  
 RESULT 9  
 GHSR\_PIG STANDARD; PRT; 366 AA.  
 AC 095254; 095255;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing peptide receptor) (GHRP) (Ghrelin receptor).  
 GN GHSR.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
 NX NCBI\_TaxId=9823;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Yorkshire; TISSUE-Pituitary;  
 RX MEDLINE=96337998; PubMed=8688086;  
 RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P., Liberato P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L., Palya O.C., Anderson J., Paress P.S., Diaz C., Chou M., Liu K.K., McKee K.K., Pong S.-S., Chung L.-Y., Elbrecht A., Dashkevitz M., Heavens R., Rigby M., Stathopoulos J., Dean D.C., Mellillo D.G., Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K., Schaefer J.M., Smith R.G., van der Ploeg L.H.T.;  
 "A receptor in pituitary and hypothalamus that functions in growth hormone release.";  
 Science 273:974-977(1996).  
 RL Science 273:974-977(1996).  
 CC -1- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins. Stimulates growth hormone secretion. Binds also other growth hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6) as well as non-peptide, low molecular weight secretagogues (e.g. L-692,429, MK-0677, adenosine).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1A (SHOWN HERE) AND 1B: ARE PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND SECRETAGOGUES.  
 CC -1- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial



```

RX MEDLINE=98344812; PubMed=9681487;
RA Sun Y.M., Millier R.P., Ho H., Gershengorn M.C., Iilling N.;
RT "Cloning and characterization of the chicken thyrotropin-releasing
RT hormone receptor.";
RL Endocrinology 139:3390-3398(1998).
CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC RECEPTOR IS MEDIATED BY G PROTEIN WHICH ACTIVATE A
CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC -----
CC DR EMBL; Y18244; CAA77091.1; -.
CC DR InterPro; IPR000276; GPCR_Rhodpsn.
CC DR Pfam; PF00001; 7tm.1; 1.
CC DR PRINTS; PR00237; GPCR_RHODOSN.
CC DR PROSITE; PS00237; G_PROTEIN_REC_P1.1; 1.
CC DR PROSITE; PS50262; G_PROTEIN_REC_P1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein.
CC FT DOMAIN 1 30
CC FT TRANSMEM 31 53
CC FT DOMAIN 54 63
CC FT TRANSMEM 64 85
CC FT DOMAIN 86 101
CC FT TRANSMEM 102 123
CC FT DOMAIN 124 146
CC FT TRANSMEM 147 170
CC FT DOMAIN 171 195
CC FT TRANSMEM 196 217
CC FT DOMAIN 218 268
CC FT TRANSMEM 269 290
CC FT DOMAIN 291 298
CC FT TRANSMEM 299 321
CC FT DOMAIN 322 395
CC FT CARBOHD 3 3
CC FT CARBOHD 10 10
CC SQ SEQUENCE 395 AA; 44697 MW; 291D9BD218723CB CRC64;
Query Match 21.4%; Score 270.5; DB 1; Length 395;
Best Local Similarity 32.0%; Pred. No. 8e-13;
Matches 72; Conservative 52; Mismatches 78; Indels 23; Gaps 6;
OY 20 OKH--LNSTEYLAFLCGPRSHFELVSVVYVIFPVGVIGNVLCVLIQHQAKTP 76
DB 9 ONHTGLISQDFV-----AEYQVVTLLVLLIGGLGIVGNIMVVLVIRTKHMRP 61
OY 77 TNYLFLAVSDLLVYL-LGMP--LEVEYEMRNRPFLFEGPVGCFKTALEFVCFASIL 132
DB 62 TNCYLVSLAVADLMVLAAGLPNTESLYKSW----YGVGGLCLITYIQGLINASSF 116
OY 133 SITTVSVRRYVALHPFAKLQSTRRLRLILGIYMGVSYLEPLNPNTISHGKIFHPNG 192
DB 117 SITAFTEIRYIALCHPIKAOFLCTFSRAKTKIIFWFSFASYCMLPFLLDLNIAVYKDT 176
OY 193 SLVGSATCIIVKPMWIMNFTIOVTSFLEPLNPVIVLYLVIA 237
DB 177 TVV-----SCGIKVSRSYSPIYMDFGIFVLPWVLAIVLYGLIA 217
RESULT 11
GHSR_MOUSE
ID GHSR_MOUSE STANDARD; PRT; 257 AA.
AC 099P50; 091282;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DT

```





```

FT TRANSMEM 62 83 2 (POTENTIAL).
FT DOMAIN 84 99 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 100 121 3 (POTENTIAL).
FT DOMAIN 122 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 168 4 (POTENTIAL).
FT DOMAIN 169 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 5 (POTENTIAL).
FT DOMAIN 216 266 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 267 288 6 (POTENTIAL).
FT DOMAIN 289 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 319 7 (POTENTIAL).
FT DOMAIN 320 393 CYTOPLASMIC (POTENTIAL).
FT CAROAMD 3 3 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CAROAMD 10 10 N-LINKED (GLCNAc. . .) (POTENTIAL).
SQ SEQUENCE 393 AA: 44559 MW: 8739B75D1A0FCCFC CRC64;

Query Match 21.0%; Score 265; DB 1; Length 393;
Best Local Similarity 28.9%; Pred. No. 2e-12;
Matches 72; Conservative 48; Mismatches 69; Indels 60; Gaps 9;

QY 15 LEDPQKHNSE-----FY-----LAFLCGPRSRHFLVSVYVPIFVGVY 58
DB 1 MENDYSEKNQTELOPQAAVALEYQVYVTLVYITG-----LGIV 41

QY 59 GAVLYCVLIQHOAMKPTNYLYFLSLAVSDLVLL-LGMPL---EYEWKRNYPFLFGV 114
DB 42 GINMVLVVMRKHKMTPTNCYIVSLAVADLVNVAAGLPNTDSTYGSN----YGVY 96

QY 115 GCYFKTALFEYVCFASISITTVSVERYVAIHPFRAKIQSTRRLRLIGYVGSYLE 174
DB 97 GCCTIYVLYQVIGLINSASSCITAFTERIYAIQHPKAOFLCFESRAKIIIFWAFSTIY 156

QY 175 SLPTNSIHGKIFHYFNPGLVPSGATC-----TVIKPMWIVYIQVSPFLYLLPMYV 228
DB 157 CHMRFLLDLNISTYKANVYV-----SCGKISKNTYSPIYLMDEGV-----FYVPMIL 206

QY 229 ISVLYYIMA 237
DB 207 ATVLYGEFIA 215

RESULT 13
TFRF_HUMAN STANDARD; PRT; 398 AA.
AC P34981;
DC 01-FEB-1994 (rel. 28, Created)
DT 01-FEB-1994 (rel. 28, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin receptor).
DE TRHR.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[11]
RP SEQUENCE FROM N.A.
RX MEDLINE=93371401; PubMed=8395824;
RA Matre V., Karlisen H.E., Wright M.S., Lundell I., Fjeldheim K., Gabrielsen O.S., Lathhammer D., Gautvik K.M.;
RT "Molecular cloning of a functional human thyrotropin-releasing hormone receptor.";
RL Biochem. Biophys. Res. Commun. 195:179-185(1993).
[12]
RP SEQUENCE FROM N.A.
RC TISSUE=Plutitary;
RX MEDLINE=93384596; PubMed=8396925;
RA Yamada M., Monden T., Satoh T., Satoh N., Murakami M., Iriuchijima T., Kakegawa T., Mori M.;
RT "Plutitary adenomas of patients with acromegaly express thyrotropin-releasing hormone receptor messenger RNA: cloning and functional expression of the human thyrotropin-releasing hormone receptor gene.";
```

```

RL Biochem. Biophys. Res. Commun. 195:737-745(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94063224; PubMed=8243797;
RA Duthie S.M., Taylor P.L., Anderson L., Cook J., Eldne K.A.;
RT "Cloning and functional characterisation of the human TRH receptor.";
RL Mol. Cell. Endocrinol. 95:R11-R15(1993).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=95002135; PubMed=7918619;
RA Hinuma S., Hosoya M., Ogi K., Tanaka H., Nagai Y., Onda H.;
RT "Molecular cloning and functional expression of a human thyrotropin-releasing hormone (TRH) receptor gene.";
RL Biochim. Biophys. Acta 1219:251-259(1994).
[5]
RP SEQUENCE FROM N.A.
RX MEDLINE=96355621; PubMed=8703031;
RA Iwasaki T., Yamada M., Satoh T., Konaka S., Ren Y., Hashimoto K., Kohga H., Kato Y., Mori M.;
RT "Genomic organization and promoter function of the human thyrotropin-releasing hormone receptor gene.";
RL J. Biol. Chem. 271:22183-22188(1996).
[6]
RP SEQUENCE OF 1-263 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=99101143; PubMed=9886052;
RA Matre V., Hovring P.T., Orstavik S., Frensen E., Rian E., Velickovic Z., Murray-McIntosh R.P., Gautvik K.M.;
RT "Structural and functional organization of the gene encoding the human thyrotropin-releasing hormone receptor.";
RL J. Neurochem. 72:40-50(1999).
-1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D16845; BA004120.1; -
DR EMBL: X75071; CAAS2965.1; -
DR EMBL: X72089; CAAS0979.1; -
DR EMBL: S75283; AAB32222.1; -
DR EMBL: S75281; AAB32222.1; JOINED.
DR EMBL: D85376; BAA12796.1; -
DR EMBL: D85375; BAA12796.1; JOINED.
DR EMBL: A0111701; CA009746.1; -.
DR PIR: JN0759; JN0759.
DR PIR: S40682; S40682.
DR PIR: JN0708; JN0708.
DR Genew: HGNC:12299; TRHR.
MIM: 188545; -.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODPSN.
DR PROSITE: PS00237; G_PROTEIN_RECPT_F1_1; 1.
DR PROSITE: PS00262; G_PROTEIN_RECPT_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Polymorphism.
FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 29 51 1 (POTENTIAL).
FT DOMAIN 52 61 2 (POTENTIAL).
FT TRANSMEM 62 83 2 (POTENTIAL).
FT DOMAIN 84 99 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 100 121 3 (POTENTIAL).
FT DOMAIN 122 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 168 4 (POTENTIAL).
FT DOMAIN 169 193 EXTRACELLULAR (POTENTIAL).
```

	FT	TRANSMEM	194	215	5 (POTENTIAL).
FT	DOMAIN	216	266		CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	267	288		6 (POTENTIAL).
FT	DOMAIN	289	296		EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	297	319		7 (POTENTIAL).
FT	DOMAIN	320	398		CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	3	3		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	10	10		N -> K (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	10	10		N -> K (IN DESNP:5774).
FT	VARIANT				/FTID-VAR-011857.
SEQUENCE	398 AA;	45084 MW;			FE920B5FE2E933D3E CRC64;
Query Match	20.5%;	Score 259;	DB 1;	Length 398;	
Best Local Similarity	32.4%;	Pred. No. 5.4e-12;			
Matches	66;	Conservative	47;	Mismatches	65;
				Indels	26;
				Gaps	
QY	45	VSVVYVPPLFV-VGYIGNVLVCILVLIQHQAQKPTNYVYFSLAVSDLLVL-LGMPL---E	99		
DB	27	VTLIELVILICGIGIVGINMVLVWRTKRMHMRPTNCYLSIALVADLMVLVLAAGLPINIDS	86		
QY	100	VYEMARNVPFLTEGPGVGYCFKTALEPFGVFASTLTSTTVSEVRYVALILHPRAKLOSTRR	159		
DB	87	IYGSN-----VGYGCLCTITLYQLVLGINASSCSTIATFTIYAIACHIKIKOICTSR	141		
QY	160	ALRLIGIVGVSFSLPNTSHTGIRKIFHPNGSLPGSATC-----TVIKPMIYNEI	213		
DB	142	AKKILIFWAFSTLCYMLMFLLDNISISPKDAIVI---SCGYKISRNYSPITLMDEG	197		
QY	214	IQVTSFLEYLLPMYIVSYLYYLA	237		
DB	198	V-----ETVVPMLIATVLYGRIA	215		

RESULT 14	ID	TRFR_BOVIN	STANDARD:	PRT:	398 AA.
AC	O46639.				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin receptor).				
DE	TRHR.				
OC	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bovinae; Bos.				
OX	NCBI_TaxID:9913;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Holstein;				
RC	MEDLINE=9815137; PubMed=9492373;				
RA	Takata M., Shimada Y., Ikeda A., Sekikawa K.;				
RT	"Molecular cloning of bovine thyrotropin-releasing hormone receptor gene".;				
RL	J. Vet. Med. Sci. 60:123-127(1998).				
CC	-I- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS				
CC	RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A				
CC	PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.				
CC	-I- SUBCELLULAR LOCATION: Integral membrane protein.				
CC	-I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.				
CC	-----				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb.ch/announce/">http://www.isb.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
CC	EMBL: D83964; BAA24069.1; -;				
CC	InterPro: IPR002767; GPCR_Rhodpsn.				
CC	Pfam: Pf00001; 7tm_1; 1.				

DR	PROSITE: PS00237; G-PROTEIN RECEPTOR_FL1; 1.
DR	PROSITE: PS0262; G-PROTEIN RECEPTOR_FL2; 1.
KW	G-protein coupled receptor; Transmembrane; Glycoprotein.
FT	DOMAIN 1 28
FT	TRANSMEM 29 51
FT	DOMAIN 52 61
FT	TRANSMEM 62 83
FT	TRANSMEM 84 99
FT	TRANSMEM 100 121
FT	DOMAIN 122 144
FT	TRANSMEM 145 168
FT	DOMAIN 169 193
FT	TRANSMEM 194 213
FT	DOMAIN 216 266
FT	TRANSMEM 267 288
FT	DOMAIN 289 297
FT	TRANSMEM 297 319
FT	DOMAIN 320 398
FT	CARBOHD 3 3
FT	CARBOHD 10 10
SEQ	SEQUENCE 398 AA: 45165 MW: D9AF4211A5701B8 CRC64:

```

Query Match Similarity      20.4%; Score 258; DB 1; Length 398;
Best Local Similarity      32.4%; Pred. No. 6.3e-12;
Matches    66; Conservative   47; Mismatches    65; Indels    26; Gaps

OY  45 VSVVVPPIFV-VGIVGNVLCLVIIIOHQA MKPTNYLFLSLVSDLLVDL-LGMPL--E 99
      ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  27 VTILLVLICGLGIIVGNIMLVLMVWRKTRKMRFPTNCIYLSSLAVADMLLVAAAGLPNTDS 86
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY  100 YEMMRNRPFLFGPRGCIFKTALFETVCFAASLIITYSVERVAIIILPFRKKLOSTRRR 159
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  87 IYGSW-----VYGVGCLCITYLYQLGINASSCSITAFIERXYIAICHPIKAQFLCTFSR 141
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY  160 ALRIIGIYWGVSFLSPTSISHGKIFHFYPNGSLVPGSGAT-----TVIKPMNIYNFI 213
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  142 AKKIIFHWAFSTICMLMFLFDLINISTYKDAIVY----SCGYKISRNYSPITLMDRG 197
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY  214 IQVTSFLFYLLPMTVISLVLYLMA 237
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  198 V-----FYVVPMLATVLYGFIA 215
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 15
TRFR-SHEEP
ID TRFR-SHEEP STANDARD: PRt: 398 AA.
AC Q28596:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberlin
DE receptor).
CN THRH..
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary.
RX MEDLINE=97200775; PubMed=9048604;
RA Boockmann J., Boeckers T.M., Winter C., Witkowski W., Winterhoff H.,
RA Deufel T., Kreutz M.R.;
RT "Thyrotropin expression in hypophyseal pars tuberalis-specific cells
RT is 3,5,3'-triiodothyronine, thyrotropin-releasing hormone, and pit-1
RT independent."
RL Endocrinology 138:1019-1028(1997).
CC -! FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -! SUBCELLULAR LOCATION: Integral membrane protein.
CC -! SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
```

```
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X95285; CAA64606.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28 1 (POTENTIAL).
FT TRANSMEM 29 51 1 (POTENTIAL).
FT DOMAIN 52 61 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 62 83 2 (POTENTIAL).
FT DOMAIN 84 99 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 100 121 3 (POTENTIAL).
FT DOMAIN 122 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 168 4 (POTENTIAL).
FT DOMAIN 169 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 5 (POTENTIAL).
FT DOMAIN 216 266 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 267 288 6 (POTENTIAL).
FT DOMAIN 289 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 319 7 (POTENTIAL).
FT DOMAIN 320 398 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 398 AA; 45088 MW; 375A311D3DD2A61A CRC64;
Query Match 20.4%; Score 258; DB 1; Length 398;
Best Local Similarity 32.4%; Pred. No. 6.3e-12;
Matches 66; Conservative 47; Mismatches 65; Indels 26; Gaps 7;
```

```
QY 45 VSVVYVPIEF-VGVIGNVLGVIIIOHQAARKPTNYVLFSLAVSDLLVIL-LGMPL---E 99
Db 27 VTLILVLIICGLGIIVINIMVLMVTKHRTTNCILVSLAVADLVAVAGLPNITDS 86
QY 100 VTEMNRNYPLEGGVGCYETKALFETVCFASISITTSVSEYRYVAILHPFRAKLOSTRR 159
Db 87 IYGSW-----VGYVGCLCTIYLYGGINASSCSITAFTERIATICHPIKAQFLCTFSR 141
QY 160 ALRIILGIWGFSLFSLPNTSIGHIKFHPNGLVPGSATC-----TVKPMIYNTI 213
Db 142 AKRIILFWAFISICYMLWFELDLNISTYKDAIV---SCGYKISRNYSPILMDFG 197
QY 214 IQVTSFLFYLLPMTVTSVLYYLMA 237
Db 198 V-----FYVPMILATVILXGPIA 215
```

Search completed: January 17, 2003, 05:04:48  
Job time : 35 secs



```

US-09-911-005-4
: Sequence 4, Application US/09911005
: Patent No. US20020119493A1
: GENERAL INFORMATION:
: APPLICANT: Glucksmann, Maria Alexandra
: TITLE OF INVENTION: 65494, A NOVEL HUMAN G-PROTEIN-COUPLED
: TITLE OF INVENTION: RECEPTOR FAMILY MEMBER AND USES THEREOF
: FILE REFERENCE: 10448-076001
: CURRENT APPLICATION NUMBER: US/09/911, 005
: CURRENT FILING DATE: 2001-07-23
: PRIOR APPLICATION NUMBER: 60/220, 042
: PRIOR FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 222
: TYPE: PR1
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: consensus sequence
: US-09-911-005-4

```

[illegible]

```

RESULT 3
US-10-080-960-25
; Sequence 25, Application US/10080960
; Publication No. US20020197695A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Glucksmann, Maria
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 80090, 52874, 52880, 63497, AND 33425
; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES THEREOF
; FILE REFERENCE: 38155-20044.00
; CURRENT APPLICATION NUMBER: US/10/080,960
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/242,040
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/242,038
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/241,992
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/242,637
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
; US-10-080-960-25

```

Query Match	20.8%	Score 262.5	DB 9	Length 259
Best Local Similarity	35.3%	Pred. No. 4.7e-18		
Matches 67	Conservative 38	Mismatches 64	Indels 21	Gaps 5

  

QY	59	GNVYLCVILIOHQAKNPTPTNYLFSVLAVSDLLVLLGLMPELEY-----EMMRNYPFELFG	112
		: :             :	
Db	1	GNLIVLILFRKKRLRPTPTNIFILMLAVADLLFLFTLPFWALYLVGSGEDMP-----FG	55
QY	113	PVGYCFKVALEFTEVFASTITTVSVSEYVVAILHPFRKLQSTR-RRLRLTIGIWMGS	171
		:         :             :	
Db	56	SALCFKVALADLVVNNKVASILLTLTASIDRYLAIYPLPKRRRTSPRRAKRYILLVWYLA	115
QY	172	VLFSLPMTSHIGIKFHPNPGSLVFGSATCTYIKP-----MMIYNEIIOVTSFLFYL	223
		:         :         :	
Db	116	LILSLPPLFLFSWVKTVESGNGTLNANNVTCLIDPEESTASVSTMILRSYVL-LSTVLGFL	174
QY	224	LPMTYISVLY 233	
		:	
Db	175	LPLVILVICY 184	

```

; RESULT 4
; US-10-012-140-27
; Sequence 27, Application US/10012140
; Publication No. US20030009017A1
; GENERAL INFORMATION:
; APPLICANT: lelby, Kevin R.
; APPLICANT: Kappel-Libermann, Rosana
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81580, AND
; TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 381552004900
; CURRENT APPLICATION NUMBER: US/10/012,140
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/246,768
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,772
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,185
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 259
; TYPE: PRT
; ORGANSIM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
; US-10-012-140-27

```

[illegible]

```

RESULT 5
US-10-012-140-29
; Sequence 29, Application US/10012140
; Publication No. US2003009017A1
; GENERAL INFORMATION:
; APPLICANT: Lelby, Kevin R.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
; TITLE OF INVENTION: 14334 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 381552004900
; CURRENT APPLICATION NUMBER: US/10/012,140
; PRIOR APPLICATION NUMBER: 60/246,768
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,772
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,185
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-10-012-140-29

Query Match 20.8%; Score 262.5; DB 9; Length 259;
Best Local Similarity 35.3%; Pred. No. 4,7e-18;
Matches 67; Conservative 38; Mismatches 64; Indels 21; Gaps 5;

OY 59 GNVLCVYLILHOAKKPTNYLFLSLAVSDLLVLLGMPLVY-----EMWRNVPFLFG 112
Db 1 GNLVILVILILKRLKRTPTNIFILNLAVADLFLTLTDPMALVYVGGSEOMP-----FG 55

OY 113 PVGCFKTALEFEVCFASILSTITVSEVRYVALHPFPAKLOSTR-RALRLILGIWGF 171
Db 56 SALCKLVYALDVNNYASILLTLTAISIDRLAYVHPLRYRRRRTSPRAKVIYLLVWVLA 115

OY 172 VLFSLPNTSHGDKRHYRPNPNSLVPGSATCYIKP-----MWIYVFIQVTSLEFYL 223
Db 116 LLTSLPPLFLFSWVKVEBGNGLNVNVTVCULIDFEEESTASYSVTLRSYVL-LSTLVGFL 174

OY 224 LPMYIVSYLV 233
Db 175 LPLVILVY 184

RESULT 6
US-10-077-698-3
; Sequence 3, Application US/10077698
; Publication No. US2003008350A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 14273 Receptor, A No. US2003008350A1el G-Protein Coupled Receptor
; FILE REFERENCE: 5800-4B, 035800/177086
; CURRENT APPLICATION NUMBER: 09/261,599
; PRIOR APPLICATION NUMBER: 09/261,599
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 09/107,761
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 09/223,538
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: patentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Unknown

```

```

: FEATURE: Description of unknown Organism: Seven
: OTHER INFORMATION: Transmembrane Segment Rhodopsin Superfamily
US-10-077-698-3

Query Match 20.8%, Score 262.5; DB 9; Length 259;
Best Local Similarity 35.3%, Pred. No. 4.7e-18;
Matches 67; Conservative 38; Mismatches 64; Indels 21; Gaps 5;

QY 59 GNVLCVLIIOHQAKTPTNYVYPSLAVSDLLVLLGMPLEY-----EMMRNYPFLFG 112
||:||||: : :|||: : :|||:|||| || :| :|
Db 1 GNLVLIIVLIRKRLRTPTNIFILNLAVADLLFLTLPPMALVYLVGSEDP-----FG 55
||:||||: : :|||: : :|||:|||| || :| :|

QY 113 PVGCYFKALEETVCFASILITTVSVERVAAILHPFRAKLOSTR-RALRLILGIWGF 171
| ||| : |||: : |||:|||| || :| :| :| :| :| :|
Db 56 SALCKLVTALDVNNMYASILLTLTAISIDRYLAIVPLRRRRRTPRRAKVVILLVWVLA 115
| ||| : |||: : |||:|||| || :| :| :| :| :| :|

QY 172 VLFSLPNTSINGIKRHFHPNGSLVPGSATCIVIKP-----MWIYFIIQVNSFLFY 223
: ||| : |||: : |||:|||| || :| :| :| :| :| :|
Db 116 LLTSLPPLFLSMVKVEEGNGTLNANNVTCCLIDPEESTASVSTWLRSYVL-LSTLVGFL 174
: ||| : |||: : |||:|||| || :| :| :| :| :| :|

QY 224 LPMFVIVSVLY 233
||: ||: ||: ||
Db 175 LPLVLIWCV 184
||: ||: ||: ||

RESULT 7
US-09-796-338A-23
: Sequence 23, Application US/09796338A
: Patent No. US20020061522A1
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc.
: TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
: FILE REFERENCE: 10448-020001
: CURRENT APPLICATION NUMBER: US/09/796,338A
: CURRENT FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: US 60/186,059
: PRIOR FILING DATE: 2000-02-29
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 23
: LENGTH: 259
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: consensus sequence
US-09-796-338A-23

Query Match 20.8%, Score 262.5; DB 10; Length 259;
Best Local Similarity 35.3%, Pred. No. 4.7e-18;
Matches 67; Conservative 38; Mismatches 64; Indels 21; Gaps 5;

QY 59 GNVLCVLIIOHQAKTPTNYVYPSLAVSDLLVLLGMPLEY-----EMMRNYPFLFG 112
||:||||: : :|||: : :|||:|||| || :| :|
Db 1 GNLVLIIVLIRKRLRTPTNIFILNLAVADLLFLTLPPMALVYLVGSEDP-----FG 55
||:||||: : :|||: : :|||:|||| || :| :|

QY 113 PVGCYFKALEETVCFASILITTVSVERVAAILHPFRAKLOSTR-RALRLILGIWGF 171
| ||| : |||: : |||:|||| || :| :| :| :| :| :|
Db 56 SALCKLVTALDVNNMYASILLTLTAISIDRYLAIVPLRRRRRTPRRAKVVILLVWVLA 115
| ||| : |||: : |||:|||| || :| :| :| :| :| :|

QY 172 VLFSLPNTSINGIKRHFHPNGSLVPGSATCIVIKP-----MWIYFIIQVNSFLFY 223
: ||| : |||: : |||:|||| || :| :| :| :| :| :|
Db 116 LLTSLPPLFLSMVKVEEGNGTLNANNVTCCLIDPEESTASVSTWLRSYVL-LSTLVGFL 174
: ||| : |||: : |||:|||| || :| :| :| :| :| :|

QY 224 LPMFVIVSVLY 233
||: ||: ||: ||
Db 175 LPLVLIWCV 184
||: ||: ||: ||

RESULT 8
US-09-823-114-20
: Sequence 20, Application US/09823114

```

```

; Patent No. US20020061554A1
; GENERAL INFORMATION:
; APPLICANT: EVANS, CHRISTOPHER J.
; KEITH, DUANE B.
; TITLE OF INVENTION: OPIOID RECEPTOR GENES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/823,114
; FILING DATE: 29-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 09/148,351
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20526.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030 MRSNFORSMH
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: group(9, 12, 33, 40, 48)
; OTHER INFORMATION: /note= "extracellular Asn residues
; that are consensus sites for N-linked glycosylation"
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-823-114-20

```

```

Query Match      20.3%; Score 256.5; DB 10; Length 415;
Best Local Similarity 28.3%; Pred. No. 36-17;
Matches 72; Conservative 49; Mismatches 90; Indels 43; Gaps 8;

QY      8 SWIYQKLDPPQKHNSDEYLAFLCGPRRSHF-----FLPVGVVYV 50
      ||:-----||:-----||:-----||:-----||:-----||:
      30 SWI-----NLSHLDGNLSDPCGPRNINLGGSDLCPTGSPSMITATTIALYS 78
      51 PLFVVGIVGNVAVCLVILLOHQAAMKPTNYLFLSLAVSDLLVLLGMPLE-VYEMWNRYP 109
      79 IYCVVGLFENFLVMYIVYTKTKTATNIIYIFNLALADALATST-LPFGSVNYLGMTPW- 136
QY      110 LGPGVGCYFKTALFETVVCASILSTTVSVERYVAIILHPFRAKLOSTRRAALILIGI 169
      ||:-----||:-----||:-----||:-----||:-----||:
      137 -FGTICKIVISIDYYNMTSIFTLCTMSVDRYIAVCHVKALDFRTPRNAKLIINCNMI 195
      170 FSVLFSLPMTSHIGIKFHFNPNGSLVPGSATCTVI--KPMWYINFLIOYTSFLF-YLLPM 226
      196 LSSAIGLPMPFMAATTKYR-----QGSIDCTLFLPSHTWENIVKICVFIFARIMPV 247
QY      227 TVISVLYLYMALRV 240
      ||:-----||:-----||:-----||:-----||:-----||:
      248 LIITVYGMILRL 261
      ||:-----||:-----||:-----||:-----||:-----||:

```

RESULT 9

```

US-09-761-962-25
; Sequence 25, Application US/09761962
; Patent No. US20020077285A1
; GENERAL INFORMATION:
; APPLICANT: Memorial Sloan-Kettering Cancer Center
; TITLE OF INVENTION: Identification and Characterization of Multiple Splice
; TITLE OF INVENTION: Variants of Mu-
; TITLE OF INVENTION: Opioid Receptor (MOR-1) Gene
; FILE REFERENCE: 830002-2000.1
; CURRENT APPLICATION NUMBER: US/09/761,962
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/743,872
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 25
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-761-962-25

```

```

Query Match      20.2%; Score 255; DB 10; Length 390;
Best Local Similarity 30.7%; Pred. No. 3,9e-17;
Matches 70; Conservative 42; Mismatches 84; Indels 32; Gaps 8;

```

```

QY      34 CGPRR-----SHFPLP-----VSVVYVPIFVVGIVGNVAVCLVILLOHQAAMKTP 76
      |||-----|||-----|||-----|||-----|||-----|||
      43 CGPNRTGLGSHSLCPQGTGSPSMVTAITIMALYSIVCVGLFENFLVMYIVYTKTKTA 102
QY      77 TNYVLFSLAVSDLLVLLGMPLE-VYEMWNRYPFLPGVGCYFKTALFETVVCASILSTIT 135
      |||-----|||-----|||-----|||-----|||-----|||
      103 TNYIYNLALADALATST-LPFGSVNYLGMTPW--FGNLTICKIVISIDYYNMTSIFTLIC 159
QY      136 TVSVERYVAIILHPFRAKLOSTRRAALILIGIYVGSVLSLPNTSHIGIKFHFNPNGSLV 195
      |||-----|||-----|||-----|||-----|||-----|||
      160 TMSVDRYIAVCHVKALDFRTPRNAKIVNWCWILSSAIGLPVPMFMAATTKYR----- 211
QY      196 PGSATCTVI--KPMWYINFLIOYTSFLF-YLLPMYISVLYLYMALRV 240
      |||-----|||-----|||-----|||-----|||-----|||
      212 QGSIDCTLFLPSHTWENILKICVFIFARIMPVLIITVYCYGMILRL 259

```

```

RESULT 10
US-09-761-962-26
; Sequence 26, Application US/09761962
; Patent No. US20020077285A1
; GENERAL INFORMATION:
; APPLICANT: Memorial Sloan-Kettering Cancer Center
; TITLE OF INVENTION: Identification and Characterization of Multiple Splice
; TITLE OF INVENTION: Variants of Mu-
; TITLE OF INVENTION: Opioid Receptor (MOR-1) Gene
; FILE REFERENCE: 830002-2000.1
; CURRENT APPLICATION NUMBER: US/09/761,962
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 09/743,872
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 26
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-761-962-26

```

```

Query Match      20.2%; Score 255; DB 10; Length 391;
Best Local Similarity 30.7%; Pred. No. 3,9e-17;
Matches 70; Conservative 42; Mismatches 84; Indels 32; Gaps 8;

```

```

QY      34 CGPRR-----SHFPLP-----VSVVYVPIFVVGIVGNVAVCLVILLOHQAAMKTP 76
      |||-----|||-----|||-----|||-----|||-----|||
      43 CGPNRTGLGSHSLCPQGTGSPSMVTAITIMALYSIVCVGLFENFLVMYIVYTKTKTA 102
QY      77 TNYVLFSLAVSDLLVLLGMPLE-VYEMWNRYPFLPGVGCYFKTALFETVVCASILSTIT 135

```



```

Db      103 TNYIENLADLALST-LPFSVNYLGTWP--FGNLTCKIVISIDYYNMTSIFTL 159
      136 TVSEVERVALHPFRKLOSTRRALRIIGIYVGFSLPNTSHGKIEHFHPNGSLV 195
      160 TMSVDRIYAVCHPVKALDFTPRNAKIVNCCNIISSAIGLPVPMFATTKYR----- 211
Qy      196 PGSATCTVI--KPMWYFNFIQVTSFLF-YLPMYVIVSVLYYLMALRV 240
      212 QGSIDCTLFHSPTWYWNELKICVFIFAFIMVLIITVCGMLIRL 259

```

# RESULT 11

US-09-761-962-19  
Sequence 19, Application US/09761962  
Patent No. US20020077285A1

```

GENERAL INFORMATION:
APPLICANT: Memorial Sloan-Kettering Cancer Center
TITLE OF INVENTION: Identification and Characterization of Multiple Splice
VARIANTS OF Mu-
FILE REFERENCE: 830002-2000.1
CURRENT APPLICATION NUMBER: US/09/761,962
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 09/743,872
PRIOR FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 19
LENGTH: 392
TYPE: PRT
ORGANISM: Mus musculus
US-09-761-962-19

```

Query Match  
Best Local Similarity 30.7%; Score 255; DB 10; Length 392;  
Pred. No. 3.9e-17;  
Matches 70; Conservative 42; Mismatches 84; Indels 32; Gaps 8;

```

Qy      34 CGRR-----SHFELP-----VSVVYVPFVGVGNVLVCLYIIOHQA MRP 76
      43 CGNRTGLGSHSLCPQTGSPSMVTAITIMALYSICVGLFNGFLMYVIVRYTKMKA 102
      77 TNYIENLADLALST-LPFSVNYLGTWP--FGNLTCKIVISIDYYNMTSIFTL 159
      103 TNYIENLADLALST-LPFSVNYLGTWP--FGNLTCKIVISIDYYNMTSIFTL 159
      136 TVSEVERVALHPFRKLOSTRRALRIIGIYVGFSLPNTSHGKIEHFHPNGSLV 195
      160 TMSVDRIYAVCHPVKALDFTPRNAKIVNCCNIISSAIGLPVPMFATTKYR----- 211
Qy      196 PGSATCTVI--KPMWYFNFIQVTSFLF-YLPMYVIVSVLYYLMALRV 240
      212 QGSIDCTLFHSPTWYWNELKICVFIFAFIMVLIITVCGMLIRL 259

```

RESULT 12  
US-09-214-904-2  
Sequence 2, Application US/09214904  
Patent No. US20010047519A1

```

GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: TRANSGENIC ANIMAL IN WHICH THE EXPRESSION
TITLE OF INVENTION: OF OPIATE RECEPTORS IS MODIFIED
NUMBER OF SEQUENCES: 6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/214,904
FILING DATE:
PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: PCT/FR97/01282
FILING DATE:
APPLICATION NUMBER: FR 96.08810
FILING DATE: 15-JUL-1996
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-214-904-2

```

Query Match  
Best Local Similarity 30.7%; Score 255; DB 10; Length 398;  
Pred. No. 4e-17;  
Matches 70; Conservative 42; Mismatches 84; Indels 32; Gaps 8;

```

Qy      34 CGRR-----SHFELP-----VSVVYVPFVGVGNVLVCLYIIOHQA MRP 76
      43 CGNRTGLGSHSLCPQTGSPSMVTAITIMALYSICVGLFNGFLMYVIVRYTKMKA 102
      77 TNYIENLADLALST-LPFSVNYLGTWP--FGNLTCKIVISIDYYNMTSIFTL 159
      103 TNYIENLADLALST-LPFSVNYLGTWP--FGNLTCKIVISIDYYNMTSIFTL 159
      136 TVSEVERVALHPFRKLOSTRRALRIIGIYVGFSLPNTSHGKIEHFHPNGSLV 195
      160 TMSVDRIYAVCHPVKALDFTPRNAKIVNCCNIISSAIGLPVPMFATTKYR----- 211
Qy      196 PGSATCTVI--KPMWYFNFIQVTSFLF-YLPMYVIVSVLYYLMALRV 240
      212 QGSIDCTLFHSPTWYWNELKICVFIFAFIMVLIITVCGMLIRL 259

```

# RESULT 13

US-09-761-962-29  
Sequence 29, Application US/09761962  
Patent No. US20020077285A1

```

GENERAL INFORMATION:
APPLICANT: Memorial Sloan-Kettering Cancer Center
TITLE OF INVENTION: Identification and Characterization of Multiple Splice
VARIANTS OF Mu-
FILE REFERENCE: 830002-2000.1
CURRENT APPLICATION NUMBER: US/09/761,962
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 09/743,872
PRIOR FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 29
LENGTH: 398
TYPE: PRT
ORGANISM: Mus musculus
US-09-761-962-29

```

Query Match  
Best Local Similarity 30.7%; Score 255; DB 10; Length 398;  
Pred. No. 4e-17;  
Matches 70; Conservative 42; Mismatches 84; Indels 32; Gaps 8;

```

Qy      34 CGRR-----SHFELP-----VSVVYVPFVGVGNVLVCLYIIOHQA MRP 76
      43 CGNRTGLGSHSLCPQTGSPSMVTAITIMALYSICVGLFNGFLMYVIVRYTKMKA 102
      77 TNYIENLADLALST-LPFSVNYLGTWP--FGNLTCKIVISIDYYNMTSIFTL 159
      103 TNYIENLADLALST-LPFSVNYLGTWP--FGNLTCKIVISIDYYNMTSIFTL 159
      136 TVSEVERVALHPFRKLOSTRRALRIIGIYVGFSLPNTSHGKIEHFHPNGSLV 195
      160 TMSVDRIYAVCHPVKALDFTPRNAKIVNCCNIISSAIGLPVPMFATTKYR----- 211
Qy      196 PGSATCTVI--KPMWYFNFIQVTSFLF-YLPMYVIVSVLYYLMALRV 240
      212 QGSIDCTLFHSPTWYWNELKICVFIFAFIMVLIITVCGMLIRL 259

```





```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S1SN5-16-A04"
/clone_lib="S1SN5"
/sex="F"
/issue_type="Ascites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-5"
/label="Top10P"
/notes="Organ: Stomach; Vector: pcNS; Site:1: EcoRI;
Site:2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deapped
with tabacco acid pyrophosphatase (TAP). The deapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT selected mRNA by
priming with dr-tailed vector. The cDNA vector was
adjusted to have about 60nl. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Bery method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

```

## BASE COUNT

```

96 a 180 c 136 g 128 t

```

## PRIMER

## Query Match

```

Best Local Similarity 72.4%; Score 528; DB 14; Length 540;
Matches 539; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

QY 3 GGAAGAACTTCAGATGCTTCCTGGATCTACAGCAGAGAACTAGAGATCCATTCCAGAA 62
Db 1 GGAAGAACTTCAGATGCTTCCTGGATCTACAGCAGAGAACTAGAGATCCATTCCAGAA 60
QY 63 ACACCTGAACAGACCGAGAGATATCGGCTTCCTCGGACCTCGCGGACCACTT 122
Db 61 ACACCTGAACAGACCGAGAGATATCGGCTTCCTCGGACCTCGCGGACCACTT 120
QY 123 CTTCCTCCCGCTGCTGCTGATGTCGCAATTTTGGTGGGGGCTCATTCGGCAATGC 182
Db 121 CTTCCTCCCGCTGCTGCTGATGTCGCAATTTTGGTGGGGGCTCATTCGGCAATGC 180
QY 183 CCGGCTGCTGCTGCTGATTCGACAGACAGGCTATGAGACGCCACCACTACTACTCT 242
Db 181 CCGGCTGCTGCTGCTGATTCGACAGACAGGCTATGAGACGCCACCACTACTACTCT 240
QY 243 CTTCAGGCTGGGCTGCTGACCTCGGCTCGGCTCTTGGAAATGCCCTGGAGGTCTA 302
Db 241 CTTCAGGCTGGGCTGCTGACCTCGGCTCGGCTCTTGGAAATGCCCTGGAGGTCTA 300
QY 303 TGAAGATGGGCGCACTACCTCTTCTGTTGCGGGCCGTTGGGCTGCTACTTCAAGACG 362
Db 301 TGAAGATGGGCGCACTACCTCTTCTGTTGCGGGCCGTTGGGCTGCTACTTCAAGACG 360
QY 363 CCGCTTTGAGACCGTGTGCTTGCCCTTCATCTCAGCATCACACCGCTGAGAGCG 422
Db 361 CCGCTTTGAGACCGTGTGCTTGCCCTTCATCTCAGCATCACACCGCTGAGAGCG 420
QY 423 CTACGCTGGCATCTACACCGCTTCGCGCCAAACTGAGAGACCGCGCGCGGCGCT 482
Db 421 CTACGCTGGCATCTACACCGCTTCGCGCCAAACTGAGAGACCGCGCGCGGCGCT 480
QY 483 CAGAGATCTCGGCATGCTCT-666GCTTCTCGGTGCTTCTTCCCTGCCAACCAGCA 541
Db 481 CAGAGATCTCGGCATGCTCTGAGGCTTCTCGGTGCTTCTTCCCTGCCAACCAGCA 540

```

## RESULT 2

## AG046172/c

```

LOCUS AG046172 670 bp DNA linear GSS 02-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-025E24.R, genomic survey sequence.

```

```

ACCESSION AG046172
VERSION AG046172.1 GI:16583064
KEYWORDS GSS.
SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
BAC library clone:PTB-025E24.R.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.

```

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

```

Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : Sacti
R.Site 2 : Sacti.
Location/Qualifiers
1..670
/organism="Pan troglodytes"
/db_xref="taxon:9606"
/clone="PTB-025E24.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC library"

```

## FEATURES

```

source
1..670

```

## BASE COUNT

```

161 a 157 c 213 g 138 t 1 others

```

## Query Match

```

Best Local Similarity 67.0%; Score 488.4; DB 17; Length 670;
Matches 492; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 231 CACTACTACTCTTTCAGCTGCGGCTCTGACCTGCTGCTGCTCTGGAATGCC 290
Db 669 CACTACTACTCTTTCAGCTGCGGCTCTGACCTGCTGCTGCTGGAATGCC 610
QY 291 CCGTAGGCTATGAGATGTGGGCGCACTACCTTTCTGTTGCGGCGCGGCTGCTA 350
Db 609 CCGTAGGCTATGAGATGTGGGCGCACTACCTTTCTGTTGCGGCGCGGCTGCTA 350
QY 351 CTTCAGACGGCCCTTTTGAACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410
Db 549 CTTCAGACGGCCCTTTTGAACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 490
QY 411 CAGCGTAGAGGCTGACGCGGCGCATCTACACCGCTTCGCGCCAAACTGAGAGACCG 470
Db 489 CAGCGTAGAGGCTGACGCGGCGCATCTACACCGCTTCGCGCCAAACTGAGAGACCG 430
QY 471 GCGCGGCGCCCTTCAGAGATCTCGGCATGCTGAGGCGCTTTCGCTGCTGCTGCTGCT 530
Db 429 GCGCGGCGCCCTTCAGAGATCTCGGCATGCTGAGGCGCTTTCGCTGCTGCTGCTGCT 370
QY 531 CAACACGAGCATGAGCATGAGCATGAGCATGAGCATGAGCATGAGCATGAGCATGAG 590
Db 369 CAACACGAGCATGAGCATGAGCATGAGCATGAGCATGAGCATGAGCATGAGCATGAG 310
QY 591 TTGCGCCACCTGTACGCTGATCAAGCCCATGATGATGATGATGATGATGATGATGAT 650
Db 309 TTGCGCCACCTGTACGCTGATCAAGCCCATGATGATGATGATGATGATGATGATGAT 250

```



D	b	143	ATGGAATAAATCTCAAAATGCTTCCTCGAGATCTACAG - AAAAACTACAAGATCCATTCCAA	201
Oy		61	AAACACCTTAACAGCACCGAGAGATATCTGGCC - TTCTCTGTGGACTCGGCAGCCA	119
D	b	202	AAAACACCTTAACAGCACCGAGAGATATCTGGCCTTCTCTGTGGACCTCGGCGAGCCA	261
Oy		120	CTTCTTCTCCCCCGTGCTCTGTGTGTATGTGCCAATTTTGTGTGTGGGCTCATTTGGCAA	179
D	b	262	CTTCTTCTCCCCCGTGCTCTGTGTGTATGTGCCAATTTATGTGTGGGCGGCTCATTTGGCAA	321
Oy		180	TGTCCGATGTCTCTGTGTATTTCTGAGACACCAGGCTATGAAGACGCCACCAACTACTA	239
D	b	322	TGTCTCGATGTCTGTGTATTTCTGAGACACCAGGCTATGAAGACGCCACCAACTACTA	380
Oy		240	CCTCTTACGCGCTGG - CGGTCTCTGACCTCTGG - TCTGTCTCTTGAATGCGCTGTGA	296
D	b	381	CCCTTTTACACCTTCGCGCGGTCTGTCTGACCTCTGTGGGTTCTTGTCTTTGGAAAGGCCCGTGG	440
Oy		297	GGTCT -- ATGAGATGTGGCGCACTACCTTTCTTGTTCGGGCC -- GTGGGCTGTCTA	350
D	b	441	GAGGTCTTATGACATGTGGCGCAACTACCTTTTGTGGTCTGGCGCCCTGTGGGCTGTAC	500
Oy		351	CTTCAAGACGCGCTCTTGTGAGACCGGTGTCTGGCTTCGCTCATCTCAGATACACACGT	410
D	b	501	ATTCAAGAGCGCGCTTTTGTGAGACCGGTGTCTGGCTTCGCTCATCTCAGATAC - CACGCT	559
Oy		411	CAGCGTGAAGCGCTACGTGGCCATCTTCACACCGCTTCGCGGCCAAACTGACAGACACCGG	470
D	b	560	CAG - GTGAGCGCGCTACGTGGCAATCTTCACACCGCTT - CGCGCAAAATTGAGAGACCGG	617
Oy		471	GCGCGCGCGCGCTCAGAGATCTCTGGCATGTCTGTGGGCGCTTCTCCGTCTTCTCCCTGGC	530
D	b	618	GGTCCCGGCGCTCAGAGATCTCTGGAATTTCTGTGGGCGCTACTCCGAGACACTTCTCCCTG --	675
Oy		531	CAACACCGACATCATGAGCATCA 553	
D	b	676	CAAAACGACATCATGAGCATCA 698	
RESULT 5				
BBS87450				
LOCUS			624 bp mRNA linear EST 26-OCT-2001	
DEFINITION			BBS87450 RIKEN full-length enriched, adult female vagina Mus	
ACCESSION			musculus cDNA clone J930001G13 5', mRNA sequence.	
VERSION			BBS87450	
KEYWORDS			BBS87450.2 GI:16449720	
SOURCE			EST.	
ORGANISM			house mouse.	
			Mus musculus	
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;	
AUTHORS			Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
			I (bases 1 to 624)	
			Arkawa,T., Carlinici,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,	
			Himoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,	
			M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,	
			Okazaki,Y., Okino,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasakih,	
			,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,	
			Tadami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,	
			Muramatsu,M. and Hayashizaki,Y.	
			RIKEN Mouse ESTs (Arkawa,T., et al. 2001)	
			Unpublished (2001)	
TITLE	JOURNAL			
COMMENT				
			On Nov 30, 2000 this sequence version replaced gi:11483994.	
			Contact: Yoshihide Hayashizaki	
			Laboratory for Genome Exploration Research Group, RIKEN Genomic	
			Sciences Center(GSC), Yokohama Institute	
			The Institute of Physical and Chemical Research (RIKEN)	
			1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, kanagawa 230-0045, Japan	
			Tel.: 81-45-503-9222	
			Fax: 81-45-503-9216	
			Email: genome-res@gsc.riken.go.jp,	
			url:http://genome.gsc.riken.go.jp/	
			Carlinici,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh	
			,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.	

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

wagii, K., Fujiiwaka, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamamaki, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences *Mamm. Genome*, 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES		Location/Qualifiers
SOURCE		1..624
		/organism="Mus musculus"
		/db_xref="taxon:10090"
		/clone="9930001G13"
		/clone_lib="RIKEN full-length enriched, adult female vagina"
		/sex="female"
		/tissue_type="vagina"
		/dev_stage="adult"
		/lab_host="DH10B"
		/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15',
		GAGGAGAGAGATCCACAGCCCTTTTTTTTTTTTTTVN 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGGAGAGATTCCTCGAGTTAATTATATCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified plucscript KS(+) after bulk excision from Lambda
		FLC I."
BASE COUNT	121 a 166 c 165 g 172 t	
ORIGIN		
Query Match	38.4%; Score 280; DB 10; Length 624;	
Best Local Similarity	77.1%; Pred. No. 9.4e-54;	
Matches 373; Conservative	0; Mismatches 95; Indels 16; Gaps 2;	
QY	1 ATGMAAAACTTCGAATGCTTCTCGATCTACCAGCAGAACTAGAAGATCCATTCAG 60	
Db	106 ATGGAAAACTTGAAMTGTCTTCGTGATCCAC-----GATTCCTCATG 150	
QY	61 AAACACCTGAAACAGACCGAGAGTATCTGGCCT-TCCTCTGCGAGACCTCGCGCAGCCA 119	
Db	151 AAGTACTTGAACACAGCAGAGAGTACTTGGCTACCCCTGTGTGGACCAAGCGCAGTGA 210	
QY	120 CTCTCTCTCCCGCTGTCTGTGGTGTATGTCCCAATTTTGTGGTGGGGGCTCAATTGCAA 179	
Db	211 CCAATCCCTCCAGATGCTGTGTGTCTATGTGGCTCATCTCTGTGGTGGGGGTATAGGCAA 270	

QY	180	TGTCCTGGGTGGCCGCGGATCTTCGACACACAGGCTATGAAAGCGCCACCACTACTA	239
Db	271	TTCTTGGGTGGCCGCGGATCTTCGACACACAGGCTATGAAAGCGCCACCACTACTA	330
QY	240	CCTCTTCAGCCCTGGGCGGCTGTGACCTCTGTGCTGCTCTCTTGGAAATGCCCTGGAGCT	299
Db	331	TCTCTTCAGCTTGGGCGGCTGTGACCTCTGTGCTGCTCTCTTGGAAATGCCCTGGAGCT	390
QY	300	CTATGAGATGTGGCGCAACTACCTCTTCTGTGTGGGCGCCGTGGGCTCTACTCAAGAC	359
Db	391	CTACGAGTGTGGCGCAACTATATCCCTCTGTGTGGGCGCGGGATGCTACTTCAAGAC	450
QY	360	GCCCTCTTTAGAGACCGGTGTGCTGTGCTCTCTCTCTCACTCAAGCATCACACCCTGAGCTGA	419
Db	451	AGCCCTTTTTCAGAGCATGTGTGCTGTGTGCTCTCTCTCTCACTCAAGCATCACACCCTGAGCTGA	510
QY	420	GCGCTACGTGGCGCATCTCTACACCCCTTCCGCGGCGCAATGTCAGACACACCGCGCGCGGC	479
Db	511	GCGCTACGTGGCGCATCTCTACACCCCTTCCGCGGCGCAATGTCAGACACACCGCGCGCGGC	570
QY	480	CCTC 483	
Db	571	CCCC 574	
RESULT 6			
LOCUS	BE751310		
DEFINITION	203132 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.		
ACCESSION	BE751310		
VERSION	BE751310.1		
KEYWORDS	EST.		
SOURCE	cow.		
ORGANISM	Bos taurus		
	Eumalpia: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
	Bovidae; Bovinae; Bos.		
REFERENCE	1 (bases 1 to 465)		
AUTHORS	Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,		
	Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett		
	G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-Mckown,C.G.,		
	Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and		
	Keane,J.W.		
TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA		
JOURNAL	libraries and construction of a gene index for cattle		
MEDLINE	Genome Res. 11 (4), 626-630 (2001)		
COMMENT	21180013		
	Contact: Smith TPL		
	USDA, ARS, US Meat Animal Research Center		
	PO Box 166, Clay Center, NE 68933-0166, USA		
	Tel: 402 762 4366		
	Fax: 402 762 4390		
	Email: smiththe@mail.marc.usda.gov		
	Single pass sequencing. Bases called and alt_trimmed with phred		
	v0.980904.e. Vector identified by cross_match with the -mismatches		
	and -mismatch 12 options.		
	PCR Primers		
	FORWARD: AGGAACAGCTATGACCAT		
	BACKWARD: GTTTCAGCTACGAGC		
	Plate: 43 row: 0 column: 6		
	Seq primer: ATTATAGTGACACTATAC.		
FEATURES			
source	1..465		
	Location/Qualifiers		

ORIGIN	BASE COUNT	88 a	129 c	113 g	135 t
Query Match	37.8%;	Score 275.4;	DB 12;	Length 465;	
Best Local Similarity	85.7%;	Pred. No. 9.5e-53;			
Matches 306;	Conservative	0;	Mismatches 51;	Indels	0;
0;					
QY	1	ATGGAAAACTTCGAATGCTCTTCCTGGATCTACACGACGAAGAACTAGAAATGCATTCACG	60		
Db	109	ATGGAAAAACATGAGAAATGTTCTTCGATGTACACACGAGAACTGAAGATCTTTTCAAG	168		
QY	61	AAACACCTGACACAGACACCGAGATATGTGCTCTCCCTCGGACCTCGGCGACGAC	120		
Db	169	AAATACCTCGAACACACACGATGACTACTACTAGTTGGCTCTCGGGGCTCTGTCGACGCCAC	228		
QY	121	TTCTTCCTCCCTCGTCTGTGGTATGTATGACCATTCTTTTGTGTGGGGGCTATGGCAAT	180		
Db	229	CTCTTCCTCCCGGATGACTGCGGTGATGCTGTGATTTTGTGTGGGGGCTGTGGCAAT	288		
QY	181	GTCCTGTGTGCTCTGTGATTTCTGACAGACACGAGCTATGAAGAGCCACCACTACTAC	240		
Db	289	CTCCGGGTGCTCTGTGATTTCTTCGCGACCAAGAGATGAAGACCCCACTTACTAC	348		
QY	241	CTCTTCACCTCGGGGGCTCTGACCTCCTCGGCTCGTCTGGAATGCCCTGGAGGTC	300		
Db	349	CTCTTCACCTTGCTGTCTTCGACCTCCTAGTCTTCTTCGGATGATCCCTTGGAAATC	408		
QY	301	TATGAGATGTGGCGCACTACCTTTCTTTTGTGGGCGCGGCTGCTACTTCAAG	357		
Db	409	TATGAGATGTGGCGCACTACCTCTCTGTTTGGGCGGCTGCTACTTCAAG	465		
RESULT 7					
B0721315		993 bp	mRNA	linear	EST 16-JUL-2002
DEFINITION	AGENCOURT 8296894	Lupsk1.symphathetic_trunk	Homo sapiens	cdna	clone
LOCUS	IMAGE:6194542	5', mRNA sequence.			
ACCESSION	B0721315				
VERSION	B0721315.1	GI:21860212			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 993)				
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a>				
	Tissue Procurement: Dr. James R. Lupski				
	CNA Library Preparation: Life Technologies, Inc.				
	CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNU)				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LLNU at:				
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>				
	Plate: LMAM13599	row: 0	column: 23		
	High quality sequence stop: 584.				
FEATURES					
SOURCE	location/qualifiers				
	1..993				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:6194542"				
	/clone_lib="Lupsk1.symphathetic_trunk"				
	/sex="male"				
	/tissue_type="sympathetic trunk"				
	/dev_stage="adult, 16 yr"				
	/lab_host="DH10B"				
	/note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:				
	NotI; Site_2: SalI; cdna made by oligo-dT priming.				
	directionally cloning using the following adaptors:				
	5'-TCGACCAACGCGTCCG-3' and				





KEYWORDS EST.  
SOURCE Japanese medaka.  
ORGANISM *Oryzias latipes*  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Perciformes; Atherinomorpha;  
Belontiiformes; Adrianiichthyidae; Oryziinae; Oryzias.  
1 (bases 1 to 624)

REFERENCE  
AUTHORS Sanaka,E., Hori,H., Naruse,K., Mitani,H. and Tanaka,M.  
TITLE Medaka EST analysis  
JOURNAL Unpublished (2001)  
COMMENT Contact: Emi Sanaka  
Department of Biological Sciences  
Graduate School of Science, Nagoya University  
Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan  
Tel: 81-52-789-2973  
Fax: 81-52-789-2974  
Email: sanaka@bio.nagoya-u.ac.jp  
This clone was isolated from Medaka eye cDNA library (SNK01) 5'end.  
Location/Qualifiers  
1. 624  
/organism="Oryzias latipes"  
/strain="wild type"  
/db\_xref="taxon:8090"  
/clone="NG147.02e"  
/clone\_lib="Medaka eye cDNA library (SNK01)"  
/tissue\_type="eye"  
/dev\_stage="adult"  
/note="Wild samples from Okayama Pref. (Southern part of Japan)"

BASE COUNT 118 a 194 c 148 g 164 t

ORIGIN

Query Match 27.5%; Score 200.6; DB 9; Length 624;  
Best Local Similarity 66.1%; Pred. No. 1.4e-35;  
Matches 290; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

121 TTCTTCCTCCCGCTGCTGTGATGTCATTTTGTGGGGGCTATGGCAAT 180  
|||||  
178 TTCTTCCTCCCGCTGACGAGGTTTACCTTCATCTTGTGATGGCCGTGTGAAAT 237  
|||||  
181 GTCTCGTGCTGCTGTGATTTGACACACAGGCTATGAGAGCGCCCAACTACTAC 240  
|||||  
238 GTCTCAACCTGTCGGGTGATGCAAAACACAGAGATGAGAACCCCAACTTCTAC 297  
|||||  
241 CTTTCAGCCTGGCGGCTGTGACTCTGCTGCTGCTCTCTTGGAATGCCCTGGAGTTC 300  
|||||  
298 CTGTGAGCCTGGCGGCTGTGACTCTGCTGATGTTGCGGATGCCCTGGAGATT 357  
|||||  
301 TATGAGATGTGGCGCACTACCCCTTCTGTTGTTGGGCCCGCTGGGCTGCTACTCAAGACG 360  
|||||  
358 TAGGACCTGTGGGAGAGACTACCGCTTCCCTTGGGAGGCGGCTGCTACTCAAGACG 417  
|||||  
361 GCCCTCTTTGAGACCGTGTGCTTGGCTCCATCTCAGACTCAACACCGCTGAGGAG 420  
|||||  
418 TTCTCTTTGAGAGCGGTGTGCTTGGCTCCATCTCAACCTCAACACCGCTGAGGAG 477  
|||||  
421 CGCTAGCGGCACTTACACCGCTTCCGCGCAAACTGACGAGCACCGGCGCGGGGCC 480  
|||||  
478 CGGTATATCGCCCTGGTGCATCTCTCAAAACCGCTTACCTGTGCACCAACGACATGCC 537  
|||||  
481 CTGAGGATCTCTGGCATCTGCTGGGCTTCTCCGCTCTTCTCCCTGGCCACACGAGC 540  
|||||  
538 AACGGGTATATACAGAGGTGTGGGTGTGATGATGTGTGCTGCTCAATCCCAACACCTCC 597  
|||||  
541 ATCCATGGCATCAAGTTCC 559  
|||||  
598 CTGCATGGCATCTTCTTAC 616  
|||||

RESULT 10  
A0244285 445 bp mRNA linear EST 15-JAN-2002  
LOCUS

DEFINITION A0244285 Medaka eye cDNA library (SNK01) *Oryzias latipes* cDNA clone  
NG147.01f, mRNA sequence.  
ACCESSION A0244285  
VERSION A0244285.1 GI:18156864  
KEYWORDS EST.  
SOURCE Japanese medaka.  
ORGANISM *Oryzias latipes*  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Perciformes; Atherinomorpha;  
Belontiiformes; Adrianiichthyidae; Oryziinae; Oryzias.  
1 (bases 1 to 445)

REFERENCE  
AUTHORS Sanaka,E., Hori,H., Naruse,K., Mitani,H. and Tanaka,M.  
TITLE Medaka EST analysis  
JOURNAL Unpublished (2001)  
COMMENT Contact: Emi Sanaka  
Department of Biological Sciences  
Graduate School of Science, Nagoya University  
Furo-cho, Chikusa-ku, Nagoya 464-8602, Japan  
Tel: 81-52-789-2973  
Fax: 81-52-789-2974  
Email: sanaka@bio.nagoya-u.ac.jp  
This clone was isolated from Medaka eye cDNA library (SNK01) 5'end.  
Location/Qualifiers  
1. 445  
/organism="Oryzias latipes"  
/strain="wild type"  
/db\_xref="taxon:8090"  
/clone="NG147.01f"  
/clone\_lib="Medaka eye cDNA library (SNK01)"  
/tissue\_type="eye"  
/dev\_stage="adult"  
/note="Wild samples from Okayama Pref. (Southern part of Japan)"

BASE COUNT 85 a 132 c 104 g 123 t 1 others

ORIGIN

Query Match 22.9%; Score 166.6; DB 9; Length 445;  
Best Local Similarity 71.0%; Pred. No. 7e-28;  
Matches 220; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

121 TTCTTCCTCCCGCTGCTGTGATGTCATTTTGTGGGGGCTATGGCAAT 180  
|||||  
136 TTCTTCCTCCCGCTGAGGTTTACCTTCATCTTGTGATGGCCGTGTGAAAT 195  
|||||  
181 GTCTCGTGCTGCTGTGATTTGACACACAGGCTATGAGAGCGCCCAACTACTACTAC 240  
|||||  
196 GTCTCAACCTGTCGGGTGATGCAAAACACAGAGATGAGAACCCCAACTTCTAC 255  
|||||  
241 CTTTCAGCCTGGCGGCTGTGACTCTGCTGCTGCTCTCTTGGAATGCCCTGGAGTTC 300  
|||||  
256 CTGTGAGCCTGGCGGCTGTGACTCTGCTGATGTTGCGGATGCCCTGGAGATT 315  
|||||  
301 TATGAGATGTGGCGCACTACCCCTTCTGTTGTTGGGCCCGCTGGGCTGCTACTCAAGACG 360  
|||||  
316 TAGGACCTGTGGGAGAGACTACCGCTTCCCTTGGGAGGCGGCTGCTACTCAAGACG 375  
|||||  
361 GCCCTCTTTGAGACCGTGTGCTTGGCTCCATCTCAGACTCAACACCGCTGAGGAG 420  
|||||  
376 TTCTCTTTGAGAGCGGTGTGCTTGGCTCCATCTCAACCTCAACACCGCTGAGGAG 435  
|||||  
421 CGCTAGCTGG 430  
|||||  
436 CGGTATATCG 445  
|||||

RESULT 11  
CNS04GXP 959 bp DNA linear GSS 21-MAY-2000  
LOCUS  
DEFINITION Tetradon nigroviridis genome survey sequence pUC-ori end of clone  
109K06 of library G from Tetradon nigroviridis; genomic survey  
sequence.  
ACCESSION AL290230

```

VERSION      AL290230.1 GI:8028810
KEYWORDS     GSS; genome survey sequence.
SOURCE       Tetraodon nigroviridis.
ORGANISM     Tetraodon nigroviridis
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
             Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
             Tetraodontidae; Tetraodon.
REFERENCE    1 (bases 1 to 959)
AUTHORS      Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
             Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
             Saurin,W. and Weissenbach,J.
TITLE        Human gene number estimate provided by genome wide analysis using
             Tetraodon nigroviridis DNA sequence
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 959)
AUTHORS      Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
             Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
             Saurin,W. and Weissenbach,J.
TITLE        Characterization and repeat analysis of the compact genome of the
             freshwater pufferfish Tetraodon nigroviridis
JOURNAL      Unpublished
REFERENCE    3 (bases 1 to 959)
AUTHORS      Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
             Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
             Saurin,W. and Weissenbach,J.
TITLE        Characterization and repeat analysis of the compact genome of the
             freshwater pufferfish Tetraodon nigroviridis
JOURNAL      Unpublished
REFERENCE    Direct Submission
AUTHORS      Submitted (12-Apr-2000)
JOURNAL      This sequence is a single read and was generated as part of a large
COMMENT      genome. For more information, please take a look at
             http://www.genoscope.cns.fr/Tetraodon.
             Location/Qualifiers
FEATURES     source
             1..959
             /organism="Tetraodon nigroviridis"
             /db_xref="taxon:99883"
             /clone="109K06"
             /note="Genoscope sequence ID : C0B6109BF03Sp1-end :
             PUC-Ori"
BASE COUNT   184 a      293 c      290 g      184 t      8 others
ORIGIN
Query Match 16.0%; Score 116.4; DB 17; Length 959;
Best Local Similarity 55.9%; Pred. No. 2.9e-16;
Matches 254; Conservative 4; Mismatches 190; Indels 6; Gaps 2;
QY 120 CTTCTCCCTCCCGTCTCTGTGATATGTCATTTTGGTGGGGGTCAATTGGCA 179
    ||||| 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 378 CTTCTGAGCGTCTCTTACAGCCTGATGCTGCYAGAGGAATCGTGGGAAACGGGTGAC 437
QY 180 TGTCTGCTGCTGCTGATTCGTGAGCAGCAGGCTATGAAGACGCCACCACTACTA 239
    || 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 438 CATCAGGTGAGCAGGAGTCTGTAGACGAGAACGGCTACTGACAGAGAGAGTCAACCA 497
240 CCTCTTACGCTGGCGTCTGTGACCTGCTGTCTGCTCTCTTGAATGCCCTGGAGGT 299
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 498 CATGATCAGCTGGCGTCTGCGACCTCTGCGTGTGCTCATCGGATCGCGGTGAGCT 557
QY 300 CTA---TAGATGTGGCGCACTACCTCTTTGTTGGGGCCCGTGGGTGCTACTTCAA 356
    || 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 558 CTACAGCGCATCTGCTTCCCTTACCTGAGGAGTCCGCAAGGTCCTGTGAAGMCTA 617
QY 357 GACGGCCCTTTTGAAGACGCTGCTTCCCTTCATCTCAGACATCACACCGTCAGCGI 416
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 618 CAACCTTCCTGTCGAGCGGTGACGCTACCGCACCATTTGAAACATAGCCAGCTTAGCTT 677
QY 417 GGAGCGCTAGTGGCCATCTACACCGTTCGCGCCCAACTGACAGGACCGCGCGCG 476
    ||||| 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 678 CGAGGCGTACGGGCGCATGTGCACACCCCTTCGCTACAAAGGCCCTGG---CGGGAAGG 734
QY 477 GGCCCTCAGAGATCTCGGAGCTGTGGGGCTTTCGCGCTCTTCTCCCTGGCCCAAC 536
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 735 CACCAAGCCCTCATCGCTTGGCGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 794

```

```

QY 537 CAGCATCCATGGCATCAAGTTCACACTTCCCC 570
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 795 CATCGCCAGCGGAGACGACGAGGCGACATCCCCCTC 828
RESULT 12
LOCUS      CNS01XJ2/c
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            sequence.
ACCESSION  AL1171767
VERSION    AL1171767.1 GI:7809824
KEYWORDS   GSS; genome survey sequence.
SOURCE     Tetraodon nigroviridis.
ORGANISM   Tetraodon nigroviridis
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
             Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
             Tetraodontidae; Tetraodon.
REFERENCE  1 (bases 1 to 817)
AUTHORS      Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
             Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
             Saurin,W. and Weissenbach,J.
TITLE        Human gene number estimate provided by genome wide analysis using
             Tetraodon nigroviridis DNA sequence
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 817)
AUTHORS      Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
             Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
             Saurin,W. and Weissenbach,J.
TITLE        Characterization and repeat analysis of the compact genome of the
             freshwater pufferfish Tetraodon nigroviridis
JOURNAL      Unpublished
REFERENCE    Direct Submission
AUTHORS      Submitted (12-Apr-2000)
JOURNAL      This sequence is a single read and was generated as part of a large
COMMENT      genome. For more information, please take a look at
             http://www.genoscope.cns.fr/Tetraodon.
             Location/Qualifiers
FEATURES     source
             1..817
             /organism="Tetraodon nigroviridis"
             /db_xref="taxon:99883"
             /clone="202F23"
             /note="Genoscope sequence ID : C0AG202CC12Sp1-end :
             PUC-Ori"
BASE COUNT   158 a      237 c      258 g      159 t      5 others
ORIGIN
Query Match 14.8%; Score 107.8; DB 17; Length 817;
Best Local Similarity 58.6%; Pred. No. 2.5e-14;
Matches 242; Conservative 1; Mismatches 163; Indels 7; Gaps 3;
QY 161 TGGTGGGGTCAATTGCAATGTCGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 220
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 481 TGGTGGGGAACGGGAGACCATCAGCGTGAAGCCAGGTCTGAAGACAGCGCTACCTGC 422
QY 221 AGACGCCCAACAACATACCTCCTCAGCCTGGCGGCTCTGCACTCCGCTGCTGCC 280
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 421 AGAAGACCTCACCGACACATGATGAGCTGACCTGCTCCACCTCTGAGTCTGCTCA 362
QY 281 TTGGAATGCCCTTGAGAGCTA---TGAGATGTGGGCACTACCTTCTTGTTCGGGC 337
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 361 TCGGATCGCGGTGAGAGCTCTACAGGCGCATCTGATCCCTTACACTGTGGCTCGGCA 302
QY 338 CGGTGGCTGCTACTTCAAGAGCGCCCTCTTGAAGACGCTGTGCTTCCCTGCATCTCA 397
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 301 ACGTCTCCGTGAAGATCTACACTTCCCTGTGAG- GCCTGAGTACGCCACCATTTGTA 243

```

OY	398	GATACCAACCCTGACCTGGAGCCCTCAGTGGCCATCTCATACCCTTCGCCGCACAAC	457
Dd	242	ACATAGCCACCGTTTAGCTTCACAGCCCTACGTGGCCATCTGCCACCCCTTCGGCTACAAG	183
OY	458	TGCAGACACCCGCGCCGCCCTTCAGAGATTCCTGCAGCATGCTGTGGGGCTTTCGGTG	517
Dd	182	CCCTGGG---CGGGGAAGCGCACCAACGGCCCTTCATCGCTTGCGCTGGTGGTGTC	126
OY	518	TTTTTCCCCCTGGCCCAACACAGCATTCATGCATCAAGTTCACACTTCCC	570
Dd	125	TGCTGGCCCTGCGCTGCTCATCGCCACGGAGACGAGGGGCACATCCCTTC	73

RESULT 13  
B603623

DEFINITION Bf603623 500 bp mRNA linear EST 25-APR-2001

ACCESSION Bf603623

VERSION Bf603623.1 GI:11701421

KEYWORDS EST.

SOURCE cow.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 500)

Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
Casas,E., Wray,J.E., White,J., Cho,J., Fahnensterg,S.C., Bennett,  
G.L., Heaton,M.P., Laegreid,W.W., Kohner,G.A., Chilko-wkown,C.G.,  
Petea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,D. and  
Keele,J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)  
21180013

COMMENT Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390

Email: smithemail.marc.usda.gov

Single pass sequencing. Bases called and alt\_trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -mismatch 12 options.

PCR Primers  
FORWARD: AGGAACAGCATATGACCAT  
BACKWARD: GTTTCOCAGTCACGACG  
Plate: 49 row: I column: 6  
Seq primer: ATTAGTGACATATAC.

FEATURES  
source Location/Qualifiers  
1..500  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC\_3BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
library made from pooled tissue from marrow, alveolar  
macrophage, ovary, fetal semitendinosus muscle, and fetal  
longissimus muscle."

BASE COUNT 71 a 179 c 139 g 111 t

ORIGIN

Query Match 14.7%; Score 107; DB 12; Length 500;  
Best Local Similarity 55.8%; Pred. No. 3.le-14;  
Matches 225; Conservative 0; Mismatches 175; Indels 3; Gaps 1;

OY 132 CGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGCGTCATGTGCATGTCTGTG 191

Dd 67 GCTCACAGCAGCAGCTCGTGGCGCTCTTGTGTGGGCAATCGCGGCACAACCTCTACCAT 126

OY 192 CTTGTGATTTGCAGACACAGCTATGAAAGCGCCACACAACTACTACTCTTAGCCT 251

[illegible]



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2003, 02:28:10 ; Search time 59 seconds

(without alignments)  
5509.118 Million cell updates/sec

Title: US-09-684-725-1

Perfect score: 729

Sequence: 1 atggaactcagaatgc.....cactcagatgagatctag 729

Index table: IDENTITY\_NUC

Gapop 10.0 , Gapect 1.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Published\_Applications\_NA.\*  
2: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/PCF\_NEM\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_NEM\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCFUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEM\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEM\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEM\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEM\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	119.8	16.4	1258	US-09-804-551B-25	Sequence 25, Appl
2	93	12.8	1427	US-09-967-768A-296	Sequence 296, App
3	89.4	12.3	1362	US-09-970-966-208	Sequence 208, App
4	89.4	12.3	1362	US-09-825-294-208	Sequence 208, App
5	85.6	11.7	1447	US-09-823-631A-12	Sequence 12, Appl
6	82.6	11.3	2108	US-09-917-800A-1620	Sequence 1620, Ap
7	82	11.2	1344	US-09-825-923-1	Sequence 1, Appl1
8	82	11.2	1344	US-09-825-923-1	Sequence 1, Appl1
9	82	11.2	1353	US-10-077-870-3	Sequence 3, Appl1
10	82	11.2	1353	US-09-825-923-3	Sequence 3, Appl1
11	79	10.8	2487	US-10-044-080-444	Sequence 444, App
12	77.8	10.7	2048	US-10-052-589-1	Sequence 1, Appl1
13	77.8	10.7	2595	US-09-964-824A-249	Sequence 249, App
14	77	10.6	2050	US-09-940-240-15	Sequence 15, Appl
15	76.8	10.5	2625	US-09-954-531-995	Sequence 995, App
16	76.8	10.5	1080	US-09-739-151-1	Sequence 1, Appl1
17	76.8	10.5	1679	US-09-104-063-5	Sequence 5, Appl1
18	75.2	10.3	1047	US-10-090-569-6	Sequence 6, Appl1
19	75.2	10.3	1805	US-09-823-114-18	Sequence 18, Appl1

20	75.2	10.3	1829	US-09-905-186A-9	Sequence 9, Appl1
21	75.2	10.3	1829	US-09-905-186A-10	Sequence 10, Appl
22	75.2	10.3	1829	US-09-905-186A-11	Sequence 11, Appl
23	75.2	10.3	1882	US-10-090-569-11	Sequence 11, Appl
24	74.8	10.3	2080	US-10-044-090-628	Sequence 628, App
25	74.8	10.3	2081	US-10-044-090-627	Sequence 627, App
26	74.8	10.3	2602	US-09-905-186A-1	Sequence 1, Appl1
27	74.4	10.2	1914	US-09-829-631A-7	Sequence 7, Appl1
28	74.4	10.2	2108	US-09-829-631A-9	Sequence 9, Appl1
29	73.8	10.1	1670	US-09-880-107-3833	Sequence 3833, Ap
30	73.8	10.1	1895	US-09-940-240-19	Sequence 19, Appl
31	73.6	10.1	1829	US-09-905-186A-7	Sequence 7, Appl1
32	73.6	10.1	1829	US-09-905-186A-8	Sequence 8, Appl1
33	73.4	10.1	720	US-10-101-487-74	Sequence 74, Appl
34	73.4	10.1	720	US-10-101-487-76	Sequence 76, Appl
35	70.6	9.7	1773	US-10-112-599A-3	Sequence 3, Appl1
36	70.4	9.7	1296	US-09-826-508-21	Sequence 21, Appl
37	70.4	9.7	1296	US-09-899-532-1	Sequence 1, Appl1
38	70.4	9.7	1710	US-09-899-532-3	Sequence 3, Appl1
39	69.6	9.5	823	US-10-044-090-598	Sequence 598, App
40	68.4	9.4	1197	US-09-796-338A-18	Sequence 18, Appl
41	68.4	9.4	1197	US-09-897-201-1	Sequence 1, Appl1
42	68.4	9.4	1609	US-09-796-338A-16	Sequence 16, Appl
43	67.2	9.2	1272	US-08-796-570A-2	Sequence 2, Appl1
44	67.2	9.2	1272	US-10-067-477-2	Sequence 2, Appl1
45	66.4	9.1	1149	US-09-993-844-10	Sequence 10, Appl

#### ALIGNMENTS

RESULT 1  
US-09-804-551B-25  
Sequence 25, Application US/09804551B  
Patent No. US2002056151A1  
GENERAL INFORMATION:  
APPLICANT: Bayer Aktiengesellschaft  
TITLE OF INVENTION: Receptors for peptides from insects  
FILE REFERENCE: ie A 34 394  
CURRENT APPLICATION NUMBER: US/09/804, 551B  
PRIOR FILING DATE: 2001-03-12  
PRIOR APPLICATION NUMBER: DE 100 13 618.4  
PRIORITY FILING DATE: 2000-03-18  
NUMBER OF SEQ ID NOS: 92  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 25  
LENGTH: 1258  
TYPE: DNA  
ORGANISM: Drosophila melanogaster  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1257)  
US-09-804-551B-25

Query Match 16.4%; Score 119.8; DB 10; Length 1258;  
Best Local Similarity 54.4%; Pred. No. 1.9e-21;  
Matches 241; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

QY	120	CTTCTCTCCCTCCGCTGCTGTGTGATGCAATTTTGTGGGGGCTATTGGCAA	179
DB	54	CATCGATGATCCGTAACGGTGTCTACCTCATTAACCGGATGAGGCA	113
QY	180	TGTCCTGCTGCTGCTGTGATTTGCAACCGCAATTTGCAACCGCAACTACTA	239
DB	114	CATAAGTACCTGATTTGATTAAGAAACCTTAATGACACGCGCAATTACTA	173
QY	240	CCCTTTACGCTGCGGCTGCTGACCTCTGCTGCTGCTGCTGGAATGCCCTTGAGGT	299
DB	174	CCCTTTTCCCTGCGCATCTCGGATTTCCCTGCTCTGTTTGTGGGGCTTCCGAGAGGT	233
QY	300	CTATGAGATGCGGCACTACCTTTTGTTCGGGCGCGGCTGCTACTTCAAGAC	359
DB	234	GTCCTACATCTGTCCAAAGTACCCGTAACGTTGGGAGTATCATCTGATCGAGCTGG	293

```
QY 360 GGCCCTTTGAGACCGTGTGCTTCGCTTCATCTCAGCATCATCCGCTCAGCGGGA 419
      |||||
Db 294 TCTGTGGGAGAGACATCGCGGATGCGACGGTGTCTACCATTAACGGCTTTCACGGTGGGA 353
QY 420 GCGCTACGTCGTCATCTTACACCGCTTCGCGGCAAACTGACAGAGACCGCGCGGGC 479
      |||||
Db 354 GCGGTATATTTGATTTGATTCGCTTTCGCGGACAGGACATGATTAACCTCACTGCGCGC 413
QY 480 CCGAGATCTCTGCGATCGTGGGGCTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 539
      |||||
Db 414 CATTCGATATGCTCTGCTTTGGATTAATGACATTAATGACCATTCGCGAGCTGCG 473
QY 540 CATTCATGTCATCAAGTTCCTACT 562
      |||||
Db 474 CCAATTTGATCGAGCATTAAT 496

RESULT 2
US-09-967-768A-296
; Sequence 296, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 296
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-296

Query Match          12.8%; Score 93; DB 10; Length 1427;
Best Local Similarity 53.5%; Pred. No. 1.4e-14;
Matches 220; Conservative 0; Mismatches 185; Indels 6; Gaps 1;

QY 143 TCGATGTGCAATTTTGTGTGGGGGTGATGGCAATGCTCTGTGTGTGTGTATTC 202
      |||||
Db 256 TCTACCGCTGGGTGCTGCTGTGGGTGTGGCAACGCCCTGTGTATCTTGTGTATTC 315
QY 203 TCGAGCACCAGGCTATGAGCGCCCAACCACTACTACTCTTACGCTGGGGTCTGTG 262
      |||||
Db 316 TCGGTACGCCCAAGATGAGAGCGGCTACCAACATCTACTGCTCAACCTGGCCGTATGGCG 375
QY 263 ACCTCTGTGCTGCTCTCTGTGAATCCCTCGAGGCTATAGATGTGGCGCACTACC 322
      |||||
Db 376 ACGAGCTCTTCACTGTGAGCGCTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCT 435
QY 323 CTTTCTTTTGGGGCGGGGGCTGTACTTCAAGAGCGCCCTCTTTGAGACGTGTGCT 382
      |||||
Db 436 C-----CTTGGGCTTCGGCTGTGTGGCGGGGTGTCAAGCTTCGAGCGGCTCAACATGT 489
QY 383 TCGCCTCATCTCTCAGCATCAACCGTGAAGGCTGAGGCTACGAGGCTATCTCAACCC 442
      |||||
Db 490 TCACAGAGCTCTTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 549
QY 443 CGTTCGCGCCCAACTGACAGACCGCGCGCGCGCGCTTCAGATCTCTGCGCATCTCT 502
      |||||
Db 550 CTCTGCTGCGCGGAGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTGT 609
QY 503 GGGGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 553
      |||||
```

```
Db 610 GGCTGSCATCCCTGTGTGTGACTCTTCCCATCGCCATCTTCCGACAGACCA 660

RESULT 3
US-09-970-966-208
; Sequence 208, Application US/09970966
; Patent No. US20020173638A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Molesh, David Alan
; APPLICANT: Fling, Steven P.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.484C6
; CURRENT APPLICATION NUMBER: US/09/970,966
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 208
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-970-966-208

Query Match          12.3%; Score 89.4; DB 9; Length 1362;
Best Local Similarity 54.4%; Pred. No. 1.1e-13;
Matches 202; Conservative 0; Mismatches 166; Indels 3; Gaps 1;

QY 164 TGGGGCTCATTTGTCATATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 223
      |||||
Db 137 TGGGGAAGGCCCCACACATTCGCTACCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 196
QY 224 GCGCCACACTACTACTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 283
      |||||
Db 197 AGAGGTGACAGACCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 256
QY 284 GAATGCCCTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 340
      |||||
Db 257 GCATGCCCTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 316
QY 341 TGGGCTGTACTTCAAGACGGGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 400
      |||||
Db 317 TGTCTGTGACGTGTGACACTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 376
QY 401 TCACACCGTGTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 460
      |||||
Db 377 TGTGTACACTGACTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 436
QY 461 AGAGCACCAGCGCGCGCGCGCTCAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 520
      |||||
Db 437 TGTGGGACCTTGTGCGAGGTGGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 496
QY 521 TCTCCCTGCCC 531
      |||||
Db 497 TGGCACTGCCC 507

RESULT 4
US-09-825-294-208
; Sequence 208, Application US/09825294
; Patent No. US20020004491A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.484C5
; CURRENT APPLICATION NUMBER: US/09/825,294
; NUMBER OF SEQ ID NOS: 215
```



Matches 278; Conservative 0; Mismatches 304; Indels 3; Gaps 1;

```
QY 126 CCTCCCGGCTGTGGTATGTCGAATTTGTGGTGGGGTCAATTTGGCAATGTCT 185
    ||| ||| ||||| | | | | | | | | | | | | | | | | | | | | |
Db 149 CATCTCTGTGGGCTGTGCTGGGCGCTTCATCTCTTGGCAATGGGCAATCTT 208
QY 186 GGTGTGCGTGTGATTCGTGACGACAGGCTATGAGAGCCGACCAACTACTCTT 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 GGTCACTCTGTGCGTGGCGCTGCAACCGGACACCTGCGAGCGCCGACCAACTACTCTT 268
QY 246 CAGCTGTGGGCTGTCTACCTCTGCTGCTCTCTCTTGGATGCCCCGAGGCTTGA 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 269 CAACCTGTGCAATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 328
QY 306 GATGTGCGGCACTACCTCTTCTGCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCT 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 AGAGTGTCTGGCTACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 385
QY 366 CTTTGAACCGTGTGCTTCCGCTTCATCTGACATGACACGCTGACGCTGAGCGCTA 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 AGATGTCTGTGCTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 445
QY 426 CGTGCCATCTTACACCGCTTCCGCGCAAACTGACAGACACCGGCGCGGCTCTAG 485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 446 CATTTGGGCTGCAATCTCTGCACTGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 505
QY 486 GATCCTGGGATGCTGCGGAGCTTCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 GCGCTCTCTGAGTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 565
QY 546 TGGCATCAAGTTCATCTTCCCAATGGGCTCCGCTGCTGCTGCTGCTGCTGCTGCT 605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 GAAAGAACCTGCGGCAATGATGACAAAGATGTGGGCTGACGAAAGAACCTCTACGC 625
QY 606 GGTTCATGAGCCCATGTGATTTATCATTCATGACAGTCACTCTCTCTCTCTCTCT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 626 CCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 685
QY 666 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 CCGGCTCTCATGCTGTGCGCAAGGACCAACGAATCTGAGGCG 730

RESULT 7
US-10-077-870-1
; Sequence 1, Application US/10077870
; Publication No. US20030003470A1
; GENERAL INFORMATION:
; APPLICANT: Salonen, Jukka T
; TITLE OF INVENTION: Method for detecting a risk of hypertension and uses thereof
; FILE REFERENCE: 0933-0183P
; CURRENT APPLICATION NUMBER: US/10/077,870
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: FI 20010323
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1341)
; OTHER INFORMATION: Coding sequence for variant human alpha-2B-adrenoceptor protein
US-10-077-870-1

Query Match 11.2%; Score 82; DB 9; Length 1344;
Best Local Similarity 52.3%; Pred. No. 8.5e-12;
Matches 206; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

QY 170 TCATTGGCAATGCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 80 TCTTGGCAACGCTGTGCTATCTCTGCTGTGTGACACGCGCTGCTGCGGCCCTC 139
QY 230 CCAACTACTACCTCTTACGCTGAGGCTCTCTACCTCTGCTGCTCTTGAATGC 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 140 AGAACCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 199
QY 290 CCGTGGAGTCTATGAGATGTGGCGGCAACTACCTCTCTGCTGCTGCTGCTGCTGCT 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 CTTTCTGCTGCGGCAACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256
QY 350 ACTTCAAGACGCGCTCTTTGAGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 AGGTGTACTGTGCGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 316
QY 410 TCAGCGTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 TCAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 376
QY 470 GCGCGCGGCGCTGAGGATCTGCGATGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 529
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 GCGGCGCATCAAGTGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
QY 530 CCAACACGACATTCATGCGCATCAAGTTCCACTA 563
Db 437 CGCCCTCATCTCAAGGCGCACAGGCGCCCA 470
```

```
RESULT 8
US-09-825-923-1
; Sequence 1, Application US/09825923
; Patent No. US2001001638A1
; GENERAL INFORMATION:
; APPLICANT: Snapir, Amir
; APPLICANT: Heinonen, Paula
; APPLICANT: Alhopuro, Pia
; APPLICANT: Karvonen, Matti
; APPLICANT: Koulu, Markku
; APPLICANT: Pesonen, Ulla-Mari
; APPLICANT: Scheinin, Mika
; APPLICANT: Salonen, Jukka T
; APPLICANT: Tuomimäen, Tomi-Pekka
; APPLICANT: Lakka, Timo A
; APPLICANT: Nyssönen, Kristina
; APPLICANT: Salonen, Riitta
; APPLICANT: Kaunonen, Jussi
; APPLICANT: Valkonen, Veli-Pekka
; TITLE OF INVENTION: A DNA molecule encoding a variant alpha-2B-adrenoceptor
; FILE REFERENCE: protein, and uses thereof
; CURRENT APPLICATION NUMBER: US/09/825,923
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/422,985
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1341)
; OTHER INFORMATION: Coding sequence for variant human
; OTHER INFORMATION: alpha-2B-adrenoceptor protein
US-09-825-923-1

Query Match 11.2%; Score 82; DB 10; Length 1344;
Best Local Similarity 52.3%; Pred. No. 8.5e-12;
Matches 206; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

QY 170 TCATTGGCAATGCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 80 TCTTGGCAACGCTGTGCTATCTCTGCTGTGTGACACGCGCTGCTGCGGCCCTC 139
```



```

QY 230 CCAACTACTACCTCTTCAAGCTGGGGCTCTGACCTCCTGGTCTGCTTTGGAATGC 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 140 AGAAGCTGTTCTGGTGTGGCTGGCGCGCGGACATCTGTTGGCAGAGCTCATCTCC 199
QY 290 CCCTGGAGGCTATGAGATGTGGCGCACTACCTTTCTTTGTTGGGCGCGTGGCTGCT 349
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 200 CTTTCTCGCTGGCCAAAGAGCTGGGCTAC---TGGTACTTCGGCGCACGTGGTGG 256
QY 350 ACTTCAAGACGGCCCTTTGAGACCGTGTGCTTGGCTCCATTCACATCACCACGG 409
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 257 AGGTGACTGGCGCTGCAGCTGCTTTCTTGACCTCTGTCATCTGACAGCTGGCGCA 316
QY 410 TCAGGCTGGAGCGCTACGTGGCGCTTACACCCGTTCCGCCCAAACTGACAGACGCC 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 TCAGGCTGGAGCGCTACGTGGCGCTGAGCGCGCTGGAGTACAACTCAAGCGACGCC 436
QY 530 CCAACACCGACATCCATGCGATCAAGTTCACCTA 563
Db 437 CCGCCCTCATCTACAAGGGCGACAGGCGCCCA 470

```

## RESULT 9

```

US-10-077-870-3
; Sequence 3, Application US/10077870
; Publication No. US20030003470A1
; GENERAL INFORMATION:
; APPLICANT: Salonen, Jukka T
; TITLE OF INVENTION: Method for detecting a risk of hypertension and uses thereof
; FILE REFERENCE: 0933-0183p
; CURRENT APPLICATION NUMBER: US/10/077,870
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: FI 20010323
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1350)
; OTHER INFORMATION: Coding sequence for human alpha-2B-adrenoceptor protein
US-10-077-870-3

```

```

Query Match          11.2% Score 82; DB 9; Length 1353;
Best Local Similarity 52.3%; Pred. No. 8.6e-12;
Matches 206; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

```

```

QY 170 TCATTGGCAATGCTGCTGTGCTGCTGATTCCTGACAGCAGCAGGCTATGAAGACGCCA 229
    ||| ||||| ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| |||
Db 80 TCTTGGCAAGCGTCTGGTCACTGCTGCTGTGACACAGCGCGTGGCGGCCCTTC 139
QY 230 CCAACTACTACCTCTTCAAGCTGGCGGCTCTGACCTCCTGCTCTGCTTTGGAATGC 289
    ||| ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 140 AGAAGCTGTTCTGTGTGCTGGCGCGCGCGCGACATCTGTTGGCGCACGTCATCATCC 199
QY 290 CCCTGGAGGCTATGAGATGTGGCGCACTACCTTTCTTTGTTGGGCGCGGCGTGTCT 349
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 200 CTTTCTCGCTGGCCAAAGAGCTGTGGCTAC---TGGTACTTCGGCGCACGTGGTGG 256
QY 350 ACTTCAAGACGGCCCTTTTGAACCGTGTGCTTGCCTTCATCTCAGCATCACCACGG 409
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 257 AGGTGACTGGCGCTGCAGCTGCTTTCTTGACACCTGTCATCGACAGCTGGCGCA 316
QY 410 TCAGGCTGGAGCGCTACGTGGCGCTTACACCCGTTCCGCCCAAACTGACAGACGCC 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 TCAGGCTGGAGCGCTACGTGGCGCTGAGCGCGCTGGAGTACAACTCAAGCGACGCC 376

```

```

QY 470 GCGCGCGCGCCCTCAGAGATCTGTGAGGCGCTTTCGCGTCTTCTCCCTGC 529
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 377 CCGCGCGCATCAAGTGCATCATCTCTACTGTGTGCTCATTCGCGCGCTCATCTCGCTGC 436
QY 530 CCAACACCGACATCCATGCGATCAAGTTCACCTA 563
Db 437 CCGCCCTCATCTACAAGGGCGACAGGCGCCCA 470

```

## RESULT 10

```

US-09-825-923-3
; Sequence 3, Application US/09825923
; Patent No. US20010016338A1
; GENERAL INFORMATION:
; APPLICANT: Snapir, Amir
; APPLICANT: Heinonen, Paula
; APPLICANT: Alhopuro, Pia
; APPLICANT: Karvonen, Matti
; APPLICANT: Koulu, Markku
; APPLICANT: Pesonen, Ullamari
; APPLICANT: Scheinin, Mika
; APPLICANT: Salonen, Jukka T
; APPLICANT: Tuomala, Tomi-Pekka
; APPLICANT: Lakka, Timo A
; APPLICANT: Nyssanen, Kristina
; APPLICANT: Salonen, Riitta
; APPLICANT: Kaunonen, Jussi
; APPLICANT: Valkonen, Veli-Pekka
; TITLE OF INVENTION: A DNA molecule encoding a variant alpha-2B-adrenoceptor
; FILE REFERENCE: protein, and uses thereof
; CURRENT APPLICATION NUMBER: US/09/825,923
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/422,985
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1350)
; OTHER INFORMATION: Coding sequence for human alpha-2B-adrenoceptor
; OTHER INFORMATION: protein
US-09-825-923-3

```

```

Query Match          11.2% Score 82; DB 10; Length 1353;
Best Local Similarity 52.3%; Pred. No. 8.6e-12;
Matches 206; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

```

```

QY 170 TCATTGGCAATGCTGCTGTGCTGATTCCTGACAGCAGCAGGCTATGAAGACGCCA 229
    ||| ||||| ||||| ||||| ||||| ||| ||| ||| ||| ||| ||| |||
Db 80 TCTTGGCAAGCGTCTGGTCACTGCTGCTGTGACACAGCGCGTGGCGGCCCTTC 139
QY 230 CCAACTACTACCTCTTCAAGCTGGCGGCTCTGACCTCCTGCTCTGCTTTGGAATGC 289
    ||| ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 140 AGAAGCTGTTCTGTGTGCTGGCGCGCGCGCGACATCTGTTGGCGCACGTCATCATCC 199
QY 290 CCCTGGAGGCTATGAGATGTGGCGCACTACCTTTCTTTGTTGGGCGCGGCGTGTCT 349
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 200 CTTTCTCGCTGGCCAAAGAGCTGTGGCTAC---TGGTACTTCGGCGCACGTGGTGG 256
QY 350 ACTTCAAGACGGCCCTTTTGAACCGTGTGCTTGCCTTCATCTCAGCATCACCACGG 409
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 257 AGGTGACTGGCGCTGCAGCTGCTTTCTTGACACCTGTCATCGACAGCTGGCGCA 316
QY 410 TCAGGCTGGAGCGCTACGTGGCGCTTACACCCGTTCCGCCCAAACTGACAGACGCC 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 TCAGGCTGGAGCGCTACGTGGCGCTGAGCGCGCTGGAGTACAACTCAAGCGACGCC 376

```

```
Qy 470 GGGCCGGGCCCCCAGAGTCCTGGGATGCTGTGGGGCTTCCTCCGCTTCCTCCCTGC 529
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 377 CGGGCGGCATCAAGTCATCATCATCTGCTGAGCTCATCGCGGCATCATCTGCTGC 436
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 530 CCAAGCACCGCATCCATGGCATCAAGTTCCACTA 563
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 437 CGCCCTCATCTACAGGCGACCAAGGCCCCCA 470
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-10-044-090-444
; Sequence 444, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 444
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Inocyte ID No. US20020137081A1 1704944CH1
US-10-044-090-444

Query Match
Best Local Similarity 10.8%; Score 79; DB 12; Length 2487;
Matches 206; Conservative 0; Mismatches 180; Indels 9; Gaps 1;

Qy 141 GGTGATGTGGCAATTTTGTGGGGGTGCATTTGGCAATGTCCTGGTGTGCTGTAT 200
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 415 GGTGAGCTGGCATTTGCTGTCTGTGCTGCAATGCTGGCAATGCTGTGGTGTGCTG 474
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 201 TCTGAGACACCAAGGCTATGAAAGACGCCACCACTACTCTTTCAGCTGGCGTCTC 260
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 475 GTGGCAACACAGCAACCTGCAGACGTACCAACTATTGTGGTGTGCTGCTGGCGGCG 534
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 261 TACCTCCCTGGTCTGCTCTTGAATGCCCTGGAGGTCTATGAGATGTGGCGCACTA 320
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 535 GCACATCGCAGTGGGTGTGCTGCGCATCCCTT-----TGCCATCACCATTAGCAA 585
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 321 CCTTTCTTTGTTGGGCGCCGTGGGCTGCTACTTCAAGACGCCCTCTTGAGACCGTGTG 380
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 586 CGGGTTTGGCGTGGCTGCGCAAGGCTGCTTTCATTGCTGCTTGTCTGTCTGCTCAC 645
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 381 CTTCGCTTCATCTCAGCATCACACGCTCAGCGTGAAGCGCTACGTTGGCCATCTTACA 440
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 646 GCAGAGCTTCATCTCAGTCTCTGCGCATCGCATTTGACCGCTTACATTTGCCATCCGAT 705
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 441 CCGGTTCGGCGGCAACATGCAAGACACCGGGCGCGCCCTCAGATCTTCGGCATCTG 500
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 706 CCGGCTCCGGTACAAATGGCTGTGTGAGACCGGCAAGAGGCTTAAGGATCATTTGCCATCTG 765
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 501 CTGGGGCTTTCCTGCTCTTCTCCCTGCCCAACA 535
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 766 CTGGGTGCTGTGCTTTGGCATTCGGCTGACTGCCA 800
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-10-052-589-1
; Sequence 1, Application US/10052589
; Patent No. US20020133832A1
; GENERAL INFORMATION:
; APPLICANT: Perez, Dianne
; APPLICANT: Zuscik, Michael
; TITLE OF INVENTION: Model systems for neurodegenerative and cardiovascular disorders
; FILE REFERENCE: 26473/04200
; CURRENT APPLICATION NUMBER: US/10/052,589
```

```
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 09/568,255
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 2048
; TYPE: DNA
; ORGANISM: Mesocricetus sp.
US-10-052-589-1

Query Match
Best Local Similarity 10.7%; Score 77.8; DB 12; Length 2048;
Matches 275; Conservative 0; Mismatches 307; Indels 3; Gaps 1;

Qy 126 CCTCCCCGTGTGTGTGTATGTGCCAATTTTGTGGTGGGGTGCATTTGGCAATGCTCT 185
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 149 CATCTCTGTGGCGTGGTGGCGCCCTTCACTCTTTGGCATTTGGGCAACATCTT 208
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 186 GGTGTGCTGGTATTCTGACAGCACAGGCTATGAAGAGCGCCACCACTACTACTCTT 245
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 209 GGTCACTCTGTAGTGGCTGTGATCGCACCTGGGAGCGCCACCACTACTACTATTGT 268
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 246 CAGCTGGCGGTCTGTGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 305
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 269 CAACCTGGCATTTGCTGACATCGCTGTGATGTTACAGTCTGCTTCTCCGCTACCT 328
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 306 GATGTGGCGGCAATACCTTTCTTTGTGGGCGCCGTGGGCTGTACTTCAAGAGCGCCCT 365
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 329 AGAGTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 385
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 366 CTTTGAGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 425
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 386 GAGCTCTCTGTGCTGTAGCGCTTCTATCTGAGCTTATGTGCACTTCTCATTTGATGCTA 445
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 426 CGTGGCATCTTACACCGCTTCCGGCGCAACCTGCAAGACACCGCGCGCGCCCTGAG 485
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 446 CATTTGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 505
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 486 GATCTCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 545
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 506 GCACTCTCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 565
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 546 TGGCATCAAGTTCCACTCTTCCCAATGAGGTCCCTGCTCCAGGTTGGCGCACTGTAC 605
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 566 GAAGAACACAGGCGCCACAGACAGAAAGATGCGGATTCACCCAGAACCTTCTATGCG 625
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 606 GGTCAATCAAGCCATGTGATCTACATTTCAATCATTCACAGTCACTTCTTCTATTCTA 665
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 626 CCTCTTTTCTCTCCGTGGGCTCTCTTACATCCACTCGGGTCAATTTGTGATGTACTG 685
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 666 CCTCTCTCCCAATGACTGTATCATGATGTCCTTACTACTCATATGCG 710
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 686 CCGGCTTACATCTGTGSCCAAGAGACCAACCAAGAACTGAGGCG 730
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-09-964-824A-249
; Sequence 249, Application us/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
```

Matches	304	Conservative	0	Mismatches	330	Indels	9	Gaps	2
QY	72	CAGCAGCAGAGATATCTGGCCTTCCTCTGGACCTCGGCGACGACCTCTTCTCC	131						
Db	168	CATCGTGTGTGACACCGCTGTGTGCCCCAACATGCCACACAAAAGCCTGTGTGTAC	227						
QY	132	CGNCTGTGTGTATGTGCAATTTTGTGTGGGGGTGATTTGGCAATGCTCGGTG	191						
Db	228	GCCTCTTATTTACATCTTTCATCTTCTGTGATCGGATGTCGCCAATCCGTGTGT	287						
QY	192	CTGTGATTTCTGAGCAGCAGGCTATGAAGACGCCACCAACTACTCTTCCACT	251						
Db	288	CTGGGTAAACATCCAGGCCAAAGACACCGGTACAGACATGCTGATCATCTTCAC	347						
QY	252	GCGCGTCTGTACCTCTCTGTCTGTCTGTGAATCCCCGTGAGGTCTATGATGTG	311						
Db	348	GGCCATCGCCACCTGTGGTGTGTGTGATCCCTGCTGTGTGTGTGTGTGTGCA	407						
QY	312	GGCAACTACCCCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA	371						
Db	408	GCATTAACAGTGGCCCATGGGGAGCTACG--TGCAAGTCAAGCACTCATCTTCTC	464						
QY	372	GACCGTGTGTCTCCCTCTCCTCAGCATCCAGCAGCTGAGGCTGAGGCTGAGC	431						
Db	465	CATCAACCTGTGTGGCAGCATCTTCTCTACGTGATGAGCTGTGACCGCTACCTC	524						
QY	432	CATCTTACACCCGTTTCCGGCCAAATGTGACAGAGACCCCGCGCCCTCAGAACT	491						
Db	525	CAT-----CACCTACTTCCGACAGCATGTGAGCGGAGAAAGGTGTGTGTGTG	578						
QY	492	CGGATGCTGTGGGGCTTCTCCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	551						
Db	579	CGTGTGTCTGT	638						
QY	552	CAATCTCACTACTTCCCAATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	611						
Db	639	GAAACCGTCAACGTGGGGGTGCTCAACAAGACCTACTGCGGCTCTTCTACCCG	698						
QY	612	CAAGCCATGTGATCTACAAATTTCAATCATCAGGTCACTCTCTATTTCACTCT	671						
Db	699	CAGGTCAAGGAGTGGGTCTACATGAGCATGTGTCTGTGTGTGTGTGTGTGTGT	758						
QY	672	CCCCATGATCTCATAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	714						
Db	759	CCCCCTGT	801						



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 03:44:22 : Search time 82 Seconds  
(without alignments)  
608.090 Million cell updates/sec

Title: US-09-684-725-2  
Perfect score: 1263  
Sequence: 1 MEKLNASWYQOKLEDPFO.....LLPMTVISVLYTLMALRWISI 242

oring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTRMBL\_21: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_virus: \*  
16: sp\_bacteriophage: \*  
17: sp\_archaeal: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	99.1	415	4	09GZQ4
2	1241	98.3	415	4	09GZQ4
3	1039.5	82.3	395	11	091276
4	1033.5	81.8	395	11	09ESQ4
5	1033.5	81.8	395	11	09JIB1
6	680.5	53.9	426	4	09HB89
7	642	50.8	402	11	09JIB2
8	642	50.8	402	11	09JIB2
9	596	47.2	405	11	055040
10	418	33.1	428	5	09VFN4
11	352	27.9	595	5	09VFN6
12	337	26.3	660	5	09VFN5
13	332.5	26.3	378	5	018701
14	303.5	24.0	418	5	017239
15	303	24.0	374	13	093412
16	301.5	23.9	363	13	093413

17	295	23.4	397	13	09DDRI	09ddr1 xenopus lae
18	290	23.0	559	13	093414	093414 spheroiodes
19	281	22.2	416	11	092005	092005 mus musculus
20	277	21.9	416	11	08VIF5	08vif5 mus musculus
21	259	20.5	401	13	09DDRO	09ddr0 xenopus lae
22	257.5	20.4	400	6	095W54	095w54 macaca fasc
23	257	20.3	464	5	09GQ54	09gq54 aedes aegypt
24	256.5	20.3	377	13	09BU14	09bu14 brachydanio
25	256	20.3	410	4	08TBH6	08tbh6 homo sapien
26	255	20.2	391	11	08VIT7	08vit7 mus musculus
27	255	20.2	391	11	08VIT7	08vit7 mus musculus
28	255	20.2	393	11	09R1M0	09r1m0 mus musculus
29	255	20.2	401	11	09R1L9	09r1l9 mus musculus
30	255	20.2	409	11	08VIF5	08vif5 mus musculus
31	255	20.2	438	11	09R0D1	09r0d1 mus musculus
32	255	20.2	444	11	09JIT1	09jit1 mus musculus
33	253.5	20.1	454	4	09H573	09h573 homo sapien
34	253	20.0	383	13	042324	042324 catostomus
35	252.5	20.0	346	4	09GGE0	09gge0 homo sapien
36	252.5	20.0	356	4	09GFE2	09gfe2 homo sapien
37	250.5	19.8	395	5	095YD7	095yd7 catostomus
38	248	19.6	404	13	09DFB0	09dfb0 catostomus
39	248	19.6	431	5	08TBD1	08tbd1 urechis uni
40	246	19.5	513	13	09DFA9	09dfa9 catostomus
41	244	19.3	370	13	08UWL5	08uwl5 fugu rubrip
42	243.5	19.3	390	13	08GQ40	08gq40 carassius a
43	243	19.2	414	6	09GK40	09gk40 canis fami
44	243	19.2	414	6	09GK99	09gk99 canis fami
45	243	19.2	443	6	09GCU1	09gcu1 canis fami

#### ALIGNMENTS

RESULT 1  
ID 09GZQ4 PRELIMINARY: PRT: 415 AA.  
AC 09GZQ4: 09NRA6:  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Neuromedin U receptor 2 (Neuromedin U receptor 2) (G  
DE protein-coupled receptor TGR-1).  
GN NMUR2 OR NMUR OR TGR-1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-20490668; PubMed-10899166;  
RA Raddatz R., Wilson A.E., Artymyshyn R., Bonini J.A., Borowsky B.,  
RA Boteju L.W., Zhou S., Kouranova E.V., Nagorny R., Guevarra M.S.,  
RA Dai M., Lerman G.S., Vaysele P.J., Branchek T.A., Gerald C., Porray C.,  
RA Adham N.:  
RT "Identification and Characterization of Two Neuromedin U Receptors  
RT Differentially Expressed in Peripheral Tissues and the Central Nervous  
RT System."  
RT J. Biol. Chem. 275:32452-32459(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX PubMed-11010960:  
RA Shan L., Qiao X., Crona J.H., Behan J., Wang S., Laz T., Bayne M.,  
RA Gustafson E.L., Konuma F.J., Jr., Hedrick J.A.:  
RT "Identification of a Novel Neuromedin U Receptor Subtype Expressed in  
RT the Central Nervous System."  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Pang L., Wang S., Laz T., Hedrick J.A.:  
RN Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 4-415 FROM N.A.

```

RX MEDLINE=20351041; PubMed=10894543;
RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
RA Zeng Z., Williams D.L., Feigheun S.D., Nunes C.N., Murphy B.,
RA Strait J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
RA Caskey T., van der Ploeg L.H.T., Liu Q.,
RT Identification of receptors for neuromedin U and its role in
RT feeding.
RL Nature 406:70-74(2000).
RN [5]
RP SEQUENCE OF 4-415 FROM N.A.
RX PubMed=10867190;
RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA Fujino M.;
RT Identification and Functional Characterization of a Novel Subtype of
RT Neuromedin U Receptor.
RL J. Biol. Chem. 275:29528-29532(2000).
DR EMBL; AF272363; AAG24794.1; -;
DR EMBL; AF292402; AAG03064.1; -;
DR EMBL; AF242874; AAF82755.1; -;
DR EMBL; AB041228; BAB13721.1; -;
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PRINTS; PR01567; NEUROMEDINUR.
DR PRINTS; PR01570; NPFRRECEPTOR.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 415 AA; 47725 MW; C2BACD84B313390F CRC64;

Query Match 99.1%; Score 1252; DB 4; Length 415;
Best Local Similarity 99.6%; Pred. No. 7e-109;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKLNQASWIIYQOKLEDPPQKHLNSTEEXYLAFLGRRSHFFLPVSVVYPIFVGVIGN 60
DB 4 MEKLNQASWIIYQOKLEDPPQKHLNSTEEXYLAFLGRRSHFFLPVSVVYPIFVGVIGN 63
QY 61 VLVCLVILQHQAKPTPTNYLFLSLAVSDLLVLLGMPLEVEYEMMRNPFLFGVGCYFKT 120
DB 64 VLVCLVILQHQAKPTPTNYLFLSLAVSDLLVLLGMPLEVEYEMMRNPFLFGVGCYFKT 123
QY 121 ALFETVCFASILSTTVSVERVYAILHPPRAKLOSTRRALRLIGIYWGFSVLSLPNTS 180
DB 124 ALFETVCFASILSTTVSVERVYAILHPPRAKLOSTRRALRLIGIYWGFSVLSLPNTS 183
QY 181 IHGIRFHYPNGLVPGSATCTVIKPMWYINFLIQVTSFLFYLLPMTVISLVLYMALRV 240
DB 184 IHGIRFHYPNGLVPGSATCTVIKPMWYINFLIQVTSFLFYLLPMTVISLVLYMALRL 243

RESULT 2
Q96AM5 PRELIMINARY; PRT; 415 AA.
AC Q96AM5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Neuromedin U receptor 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016938; AAH16938.1; -;
DR InterPro: IPR000276; GPCR_Rhodopsn.

```

```

DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PRINTS; PR01567; NEUROMEDINUR.
DR PRINTS; PR01570; NPFRRECEPTOR.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 415 AA; 47770 MW; 30BPEDD706436AB9 CRC64;

Query Match 98.3%; Score 1241; DB 4; Length 415;
Best Local Similarity 99.2%; Pred. No. 7.5e-108;
Matches 238; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MEKLNQASWIIYQOKLEDPPQKHLNSTEEXYLAFLGRRSHFFLPVSVVYPIFVGVIGN 60
DB 4 MEKLNQASWIIYQOKLEDPPQKHLNSTEEXYLAFLGRRSHFFLPVSVVYPIFVGVIGN 63
QY 61 VLVCLVILQHQAKPTPTNYLFLSLAVSDLLVLLGMPLEVEYEMMRNPFLFGVGCYFKT 120
DB 64 VLVCLVILQHQAKPTPTNYLFLSLAVSDLLVLLGMPLEVEYEMMRNPFLFGVGCYFKT 123
QY 121 ALFETVCFASILSTTVSVERVYAILHPPRAKLOSTRRALRLIGIYWGFSVLSLPNTS 180
DB 124 ALFETVCFASILSTTVSVERVYAILHPPRAKLOSTRRALRLIGIYWGFSVLSLPNTS 183
QY 181 IHGIRFHYPNGLVPGSATCTVIKPMWYINFLIQVTSFLFYLLPMTVISLVLYMALRV 240
DB 184 IHGIRFHYPNGLVPGSATCTVIKPMWYINFLIQVTSFLFYLLPMTVISLVLYMALRL 243

RESULT 3
Q91Z76 PRELIMINARY; PRT; 395 AA.
AC Q91Z76;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Neuromedin U receptor type 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Funes S., Hedrick J.A., Yang S., Shan L., Bayne M., Monsma F.J. Jr.,
RA Gustafson E.L.;
RT "Characterization of murine neuromedin U R2 receptor."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY057384; AAL26695.1; -;
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PRINTS; PR01567; NEUROMEDINUR.
DR PRINTS; PR01570; NPFRRECEPTOR.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 395 AA; 44844 MW; 0BB540024566903D CRC64;

Query Match 82.3%; Score 1039.5; DB 11; Length 395;
Best Local Similarity 81.7%; Pred. No. 4.5e-89;
Matches 196; Conservative 19; Mismatches 20; Indels 5; Gaps 1;

QY 1 MEKLNQASWIIYQOKLEDPPQKHLNSTEEXYLAFLGRRSHFFLPVSVVYPIFVGVIGN 60
DB 1 MCKLENAWIIH-----DSLMTKYNSTEEYLAFLGCRKSDLSPLSVYVALLFVGVIGN 55
QY 61 VLVCLVILQHQAKPTPTNYLFLSLAVSDLLVLLGMPLEVEYEMMRNPFLFGVGCYFKT 120
DB 56 VLVCLVILQHQAKPTPTNYLFLSLAVSDLLVLLGMPLEVEYEMMRNPFLFGVGCYFKT 115
QY 121 ALFETVCFASILSTTVSVERVYAILHPPRAKLOSTRRALRLIGIYWGFSVLSLPNTS 180

```

```

|||||
Db 116 ALPEYOCFASILSVTVTSIERVVAIVHPFRAKLESTRRRRLRLISLWSVSVFSLPNTS 175
OY 181 IHGKIFHPNGSLVPGSATCTVYIKPMWYNFIIOVTSFLFLLPMTVISLVLYMALRV 240
Db 176 IHGKIFQOPNGSSVPSGATCTVTKPMWYNFIIOATSFLLFLLPMTLISLVLYLMGLRL 235

RESULT 4
O9ES04 PRELIMINARY; PRT; 395 AA.
ID O9ES04
AC O9ES04;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
G protein-coupled receptor TGR-1.
TGR-1.
GN
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20449029; PubMed=10887190;
RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA Fujino M.,
RT "Identification and Functional Characterization of a Novel Subtype of
RT Neuromedin U Receptor."
J. Biol. Chem. 275:29528-29532(2000).
DR EMBL: AB041229; BAB13722.1;
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR pfam: PF00001; 7tm_1.1;
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PRINTS; PR01567; NEUROMEDINUR.
DR PRINTS; PR01570; NPFRRECEPTOR.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1.1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1.2; 1.
KW Receptor.
SQ SEQUENCE 395 AA; 44756 MW; EABD6D36E9C355DA CRC64;

Query Match 81.8%; Score 1033.5; DB 11; Length 395;
Best Local Similarity 80.4%; Pred. No. 1.6e-88;
Matches 193; Conservative 22; Mismatches 20; Indels 5; Gaps 1;

OY 1 MELONASWYIOQKLEDPFOKHLNSTEYLAFLCGPRRSHFLPVSVYVPIFVGVIGN 60
Db 1 MGLKLNASWTH-----DPLMKYLNSTEEYLAHLGPKRSDLSIPVSVAVALIFLVGMGN 55

OY 61 VLVCIVLIHQAMKTPNTYLFSLAVSDLLVLLGMPLEYEMMRNPFLFGVGCYFKT 120
Db 56 LVCVMYIVRHQTLKTPNTYLFSLAVSDLLVLLGMPLEYEMMRNPFLFGVGCYFKT 115

OY 121 ALPEYOCFASILSVTVTSIERVVAIVHPFRAKLOSTRRRRLRLISLWSVSVFSLPNTS 180
Db 116 ALPEYOCFASILSVTVTSIERVVAIVHPFRAKLESTRRRRLRLISLWSVSVFSLPNTS 175

OY 181 IHGKIFHPNGSLVPGSATCTVYIKPMWYNFIIOVTSFLFLLPMTVISLVLYMALRV 240
Db 176 IHGKIFQOPNGSSVPSGATCTVTKPMWYNFIIOATSFLLFLLPMTLISLVLYLMGLRL 235

RESULT 5
O9JIB1 PRELIMINARY; PRT; 395 AA.
ID O9JIB1
AC O9JIB1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
Neuromedin U receptor 2.
NMU2R.
GN
OS Rattus norvegicus (Rat).

```

```

OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=20351041; PubMed=10894543;
RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
RA Zeng Z., Williams D.L., Feiguer S.D., Nunes C.N., Murphy B.,
RA Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
RA Hrenluk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
RA Caskey T., van der Ploeg L.H.T., Liu Q.;
RT "Identification of receptors for neuromedin U and its role in
RT feeding."
RL Nature 406:70-74(2000).
DR EMBL: AF242875; AAF82756.1;
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR pfam: PF00001; 7tm_1.1;
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PRINTS; PR01565; NEUROMEDINUR.
DR PRINTS; PR01567; NEUROMEDINUR.
DR PRINTS; PR01570; NPFRRECEPTOR.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1.1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1.2; 1.
KW Receptor.
SQ SEQUENCE 395 AA; 44722 MW; 0TD3765B5D5355C0 CRC64;

Query Match 81.8%; Score 1033.5; DB 11; Length 395;
Best Local Similarity 80.4%; Pred. No. 1.6e-88;
Matches 193; Conservative 22; Mismatches 20; Indels 5; Gaps 1;

OY 1 MELONASWYIOQKLEDPFOKHLNSTEYLAFLCGPRRSHFLPVSVYVPIFVGVIGN 60
Db 1 MGLKLNASWTH-----DPLMKYLNSTEEYLAHLGPKRSDLSIPVSVAVALIFLVGMGN 55

OY 61 VLVCIVLIHQAMKTPNTYLFSLAVSDLLVLLGMPLEYEMMRNPFLFGVGCYFKT 120
Db 56 LVCVMYIVRHQTLKTPNTYLFSLAVSDLLVLLGMPLEYEMMRNPFLFGVGCYFKT 115

OY 121 ALPEYOCFASILSVTVTSIERVVAIVHPFRAKLOSTRRRRLRLISLWSVSVFSLPNTS 180
Db 116 ALPEYOCFASILSVTVTSIERVVAIVHPFRAKLESTRRRRLRLISLWSVSVFSLPNTS 175

OY 181 IHGKIFHPNGSLVPGSATCTVYIKPMWYNFIIOVTSFLFLLPMTVISLVLYMALRV 240
Db 176 IHGKIFQOPNGSSVPSGATCTVTKPMWYNFIIOATSFLLFLLPMTLISLVLYLMGLRL 235

RESULT 6
O9HB89 PRELIMINARY; PRT; 426 AA.
ID O9HB89
AC O9HB89; 043664;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
Neuromedin U receptor 1 (Orphan G protein-coupled receptor).
NMUR1.
GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20490668; PubMed=10899166;
RA Raddatz R., Wilson A.E., Artymyshyn R., Bonini J.A., Borowsky B.,
RA Boteju L.W., Zhou S., Kouranova E.V., Nagorny R., Guevarra M.S.,
RA Dai M., Lerman G.S., Vayese P.J., Branchek T.A., Gerald C., Forray C.,
RA Adham N.;
RT "Identification and Characterization of Two Neuromedin U Receptors
RT Differentially Expressed in Peripheral Tissues and the Central Nervous
RT System."
J. Biol. Chem. 275:32452-32459(2000).
RN [2]

```

Query Match	53.9%;	Score 680.5;	DB 4;	Length 426;
Best Local Similarity	56.0%;	Pred. No. 1.5e-55;		
Matches 126;	Conservative 42;	Mismatches 54;	Indels 3;	Gaps 2

214 PDSAVCM LVRPRALYNMVQTALLFCCLPMAIMSVLYLLIGRL 258

Pfam; PF00001; 7tm\_1; 1.  
PRINTS; PR00237: GPCR RHODOPSIN.

R PRINTS; PRO1565; NEUROMEDINTR.  
 R PRINTS; PRO1566; NEUROMEDINTR.  
 R POSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; UNKNOWN\_1.  
 R POSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 R Receptor.  
 O SEQUENCE 402 AA; 4578 MW; CB576083D9F7A095 CRC64

Query Match	50.8%;	Score 642;	DB 11;	Length 402;
Best Local Similarity	54.1%;	Pred. No. 5.7e-52;		
Matches 125;	Conservative 36;	Mismatches 64;	Indels 6;	Gaps 2

185 PCRGPVPDSVTCTLVRPQFFYKLVITQTTILLFFCLPMVTVISVLYLLGLRL 235

	SEQUENCE FROM N.A.	
W	MEDLINE=20347213; Pubmed=10783389;	
X	Fuji R., Hosoya M., Fukusumi S., Kawamata Y., Habata Y., Hinuma S.,	
A	Onda H., Nishimura O., Fujino M.;	
T	"Identification of neuropeptide u as the cognate ligand of the orphan G	
L	protein-coupled receptor FM-3 ";	
I	J. Biol. Chem. 275:21068-21074(2000).	
R	EMBL: AB038649; BAA9387.1;-	
R	InterPro: IPR00276; GPCR_Rhodopsn.	
R	Pfam: PF00001; 7tm_1, 1.	
R	PRINTS: PR00237; GPCRRHODOPSN.	
R	PRINTS: PR01565; NEUROMEDINUR.	
R	PRINTS: PR01566; NEUROMEDINUR.	
R	PROSITE: PS00237; G_PROTEIN_RECPT_P1_1; UNKNOWN_1.	
R	PROSITE: PS50262; G_PROTEIN_RECPT_P1_2; 1.	
R	RECEPTOR.	
D	SEQUENCE	412 AA; 46784 MW; 2CAB56299E2EA7B8 CRC64;
C	09JUI5; PRELIMINARY; PRT; 412 AA.	
G	09JUI5; 01-OCT-2000 (TREMBLrel. 15, Created)	
T	01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
T	01-JUN-2002 (TREMBLrel. 21, Last annotation update)	
N	FM-3. G protein-coupled receptor FM-3.	
S	Rattus norvegicus (Rat).	
S	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
C	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
X	NCBI_TaxID=10116;	
N	[1]	

```

Query Match      50.8%  Score 642;  DB 11;  Length 412;
Best Local Similarity  54.1%;  Pred. No. 5,8e-52;
Matches 125;  Conservative 36;  Mismatches 64;  Indels 6;  Gaps 2;

16  EDPPKKH-----LNSTEEYLAFLCGPRSRHFFLPSPVYVYVIFVWGVGANGVLCVILQ 69
   :  : : :  : : : :  : : : :  : : : :  : : : :  : : : :  : : : :
15  DSEFKKHEDLEDLNLTHERLRLKYGPPQYKQFPLPCVITYLLIFVWGVGANGITCVILR 74
   :  : : :  : : : :  : : : :  : : : :  : : : :  : : : :  : : : :

70  HOAMKPIPIYYFLSLAVSDSLVLLGMPLEVEYEMMRNRPFLGPGVCFKTALETFECPA 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
75  QKAMHPTNPFYFLSLAVSDSLVLLVGLPFLPLIEMQHNYPFQIAGAGCYRILLLEFYVCLA 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

130 SLISLTTYSVERVYVAILHDFRAKLQSTRRRLRLIGIYWGVSFLSPNTSIHGKIFHYE 189

```



```

Db 135 SVLNTALSVERYAAVHVPLQAKSVMTRTVHVRRLGAIVFAILESLPTSLHGLSPLYV 194
QY 190 PNGSLVGSATCTVYIKPMYINFIQVTSFLFYLLPMTVSVLYYLMALV 240
Db 195 PCRGVPDVSCTLVPRQPFYKLVITQTLFFCLPMVTISVLYLLGLRL 245

RESULT 9
055040 PRELIMINARY; PRT; 405 AA.
AC 055040;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, last sequence update)
DB 01-JUN-2002 (Tremblrel. 21, last annotation update)
ORphan G protein-coupled receptor.
GPR66.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RA McKee K.K., Tan C.T., Liu J., Palyna O.C., Feighner S.D.,
RA Hreniuk D.L., Smith R.G., Howard A.D.,
RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF044602; AAC02681.1; -.
DR MGD: MGI:1341898; Gpr66.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSN.
DR PRINTS: PR01565; NEURONMEDIUM.
DR PRINTS: PR01566; NEURONMEDIUM.
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR_FL_1; UNKNOWN_1.
DR PROSITE: PS50262; G-PROTEIN_RECEPTOR_FL_2; 1.
DR KEGG:
DR Receptor.
SQ SSQUNCE 405 AA; 45609 MW; F1BA493D3EB81F34 CRC64;

Query Match 47.2%; Score 596; DB 11; Length 405;
Best Local Similarity 50.6%; Pred. No. 1,1e-47;
Matches 118; Conservative 40; Mismatches 67; Indels 8; Gaps 3;

QY 9 WIIYQKLEDFQKHLNSTEYLAF-LGPRRSHEFLPVSVVYVPIFVGVIGNVLCVI 67
Db 10 WYQYR-ED-----LNLTDALRLKYLGPQOMQFVPICTYLLIFVGLGNGLTCTVI 62
QY 68 LQHQAMTPTNYIFSLAVSDLVLLGLMPELYEMKRNTPFLFGVGCYFKTALFETVC 127
Db 63 LRNTMTPTNPFYLFSLAVSDMLVLVGLPLEYEMQONPFOGASACYFRILLETVC 122
QY 128 FASLSTTTSVSVRYVALILHPRAKLQSTRRALRIILGIYWGFSVLSLPTSHGIFKH 187
Db 123 LASTLVNTALSVERYAAVHVPLQAKSVMTRAHVRMGAIVLTLFSLPTSLHGLSOL 182
QY 188 YFPGSLVPGSATCTVYIKPMYINFIQVTSFLFYLLPMTVSVLYYLMALV 240
Db 183 TVPCRGVPDVSATCTVYIKPMYINFIQVTSFLFYLLPMTVSVLYYLMALV 245

RESULT 10
055040 PRELIMINARY; PRT; 428 AA.
AC 055040;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DB 01-JUN-2002 (Tremblrel. 21, last annotation update)
DB CG9318 protein.
CG9318.
Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Preygotia; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;

```

```

RN [1]
SEQUENCE FROM N.A.
RC SRRAIN-BERKELEY;
RP MEDLINE-20196006; PubMed-10731132;
RX Adams M.D., Ceinalker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amandlides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sultion G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek J.R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Adair J.F., Agdayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson R.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshirel A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacble J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith T.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Teclor C., Turner R., Venier E., Wang A., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195(2000).
CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL: AEO03703; AAF55016.1; -.
DR FlyBase: FBgn0038201; CG9318.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSN.
DR PRINTS: PR01565; NEURONMEDIUM.
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR_FL_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECEPTOR_FL_2; 1.
DR KEGG:
DR G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ SSQUNCE 428 AA; 47198 MW; 9DDA41E93046BD7 CRC64;

Query Match 33.1%; Score 418; DB 5; Length 428;
Best Local Similarity 40.9%; Pred. No. 4.8e-31;
Matches 85; Conservative 49; Mismatches 62; Indels 12; Gaps 4;

QY 35 GPRRS--HFLPVSVVYVPIFVGVIGNVLCVIILQHQAMKPTNYIFSLAVSDLV 92
Db 11 GPRDPLAIVIPVTVVSLFIFGVGNISPCYIKKRNHNTFNYIFSLAISDFLL 70
QY 93 LCGHPELYEMKRNTPFLFGVGCYFKTALFETVCASISITTSVSVRYVALILHPRAK 152
Db 71 LSGVPEVSYIWSKYPVFGEYICIGGLAETSANATVLTITAFYERYAICHPGLQ 130
QY 153 LOSTRRRALRIILGIYWGFSVLSLPTSHGIFKHPRNSLIVGSGSATCTVYIKPMYIN 212
Db 131 AMSKLSAIRIYVWIMATVTAIPQAAFGIE-HY-----SGVDCGIVRYVHNSF 182
QY 213 IIOVTSFLFYLLPMTVSVLYYLMALV 240

```



RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287.2185-2195(2000).  
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 DR EMBL: AE003699; AAF54930.1;  
 DR FlyBase: FBgn0038140; CG8784.  
 DR InterPro: IPR000923; BlueCu\_1.  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS: PR00237; GPCRHHODPSN.  
 DR PRINTS: PR01565; NEURROMEDINUR.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 DR PROSITE: PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 DR G-protein coupled receptor; Glycoprotein; Transmembrane.  
 KW PROSITE: PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 SO SEQUENCE 660 AA; 72277 MW; C52D1EA7E4ACD39 CRC64;

Query Match 26.7%; Score 337; DB 5; Length 660;  
 Best Local Similarity 34.1%; Pred. No. 2.6e-23;

Matches 84; Conservative 42; Mismatches 86; Indels 34; Gaps 7;

QY 17 DPKQH---LNSTEEYLAFLCG-----PRSRHFFL--PVSVVYVPFVVG 56  
 DB 62 DKFLTHVAHLNITTEMLSLGSTNGNASTMAADSPVDESILTRALVICALIFVAG 121  
 QY 57 VIGNVLCVLILOHQAOKPTNYLFLSLAVSDLLVLLGMPLEVEYEMRNYP--FLGFPV 114  
 DB 122 VIGNVLCVLILOHQAOKPTNYLFLSLAVSDLLVLLGMPLEVEYEMRNYP--FLGFPV 179  
 QY 115 GCYEKTALEFVFCFASISITTVSVERVYALHPFRAKLOSTRBRALRIIGIWFGEVL 174  
 DB 180 MCIMGSLVLSMANATVLTITTAFTVERVYALHPFRAKLOSTRBRALRIIGIWFGEVL 239  
 QY 175 SLPTSHIGIKFHPFGSLVPGSATCTVIRKPMIYNIQVTSFLFYLLPMTVIVSLY 234  
 DB 240 ALPQAMFVSVYQ-----NEGVSCTMEND--FYAHVAVSGFIPEGPMALICLVY 269  
 QY 235 LMLRV 240  
 DB 290 LIGVKL 295

RESULT 13

ID 018701 PRELIMINARY; PRT; 378 AA.  
 AC 018701;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE C48C5.1 protein.  
 GN C48C5.1.  
 OS *Caenorhabditis elegans*.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; *Caenorhabditis*.  
 OC NCBI\_TaxID=6239;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Alnscoough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurtry A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,  
 RA Smaison N., Smith A., Sonhammer E., Staden K., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Wellstock L., Wilkinson-Sproat J., Wohlmann P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C.*

RT *elegans*.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Favello A.;  
 RT "The sequence of *C. elegans* cosmid C48C5.";  
 RT Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U39994; AAB37017.1;  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PRINTS: PR00237; GPCRHHODPSN.  
 DR PRINTS: PR01565; NEURROMEDINUR.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECEP\_F1\_1.  
 DR PROSITE: PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 SO SEQUENCE 378 AA; 43631 MW; 45BDC60CCBF88F86 CRC64;

Query Match 26.3%; Score 332.5; DB 5; Length 378;  
 Best Local Similarity 33.6%; Pred. No. 4.1e-23;

Matches 87; Conservative 37; Mismatches 98; Indels 37; Gaps 6;

QY 20 QKHLNSTEEY---LAFICGRRSH-----FLLPVSVVYVPF 53  
 DB 3 QACINTTEDODCCIAFCPTVYSISESEKACYEHCFISKRALDVTLYKVALYFIF 62  
 QY 54 VVGVIGVLCVLILOHQAOKPTNYLFLSLAVSDLLVLLGMPLEVEYEMRNYPFLGFP 113  
 DB 63 LVGVIGVLCVLILOHQAOKPTNYLFLSLAVSDLLVLLGMPLEVEYEMRNYPFLGFP 122  
 QY 114 GCYEKTALEFVFCFASISITTVSVERVYALHPFRAKLOSTRBRALRIIGIWFGEV 172  
 DB 123 YIGNKALIAETSSVILILFALIERVAVVCHPLEFMKVOPEKRNIGITFTWFSI 182  
 QY 173 LFLSPNTHIGIKF--HYFP---NGSLVPGSATCTV-----KPMIYNIQVTSFLY 222  
 DB 183 LCAMPFAIHRADYIMKSWGTDNRIPVSKSCMCAVMEPEKLASTKILFHPSAIAFF 242  
 QY 223 LLPMTVIVSLYLMALRVS 241  
 DB 243 ALPLFTIIVLIARIACKVS 261

RESULT 14

ID 017239 PRELIMINARY; PRT; 418 AA.  
 AC 017239;  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Hypothetical 47.7 kDa protein.  
 GN K10B4.4.  
 OS *Caenorhabditis elegans*.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; *Caenorhabditis*.  
 OC NCBI\_TaxID=6239;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RA "Genome sequence of the nematode *C. elegans*: a platform for  
 RT investigating biology. The *C. elegans* Sequencing Consortium.";  
 RL Science 282:2012-2016(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Sammons L., Wohlmann P., Antoniou B.,  
 RT "The sequence of *C. elegans* cosmid K10B4.";

```

RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF025463; AAB71009.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR Hypothetical protein.
SQ SEQUENCE 418 AA; 47735 MW; CE0416539CA3BB27 CRC64;

Query Match 24.0%; Score 303.5; DB 5; Length 418;
Best Local Similarity 30.2%; Pred. No. 2.3e-20;
Matches 76; Conservative 54; Mismatches 89; Indels 33; Gaps 6;

QY 22 HLNTEEYLAFLGPR--RSHFLPVSVYVPIFVGVIGNVLCVLIQHQAARKTPTNY 79
DB 9 NWSITITVSTLGEKQSGAIVPIYITGLFLGLFGNICTCIYIAANKSMHNPNTY 68
80 YLFSLAASDLVLLGLMPLVEYE--MMRNYPFLFGPVGCYFETKALFETVCFASILSTTVS 138
DB 69 YLFSLAASDLIALLGLMPEFYQSLDYSYPRSEGLCKARAFLETTSTASIMILCCFS 128
QY 139 VERYVALHPRAKLQSTRRALRLILGIWGFVSLFSLPNTSINGI----- 184
DB 129 FERRLAICHPLRSKIFSTLRANVLLIAMIISFVCAPIAFIYQINKPLPEDAKYQPM 188
QY 185 --KHYFRP-----NGSLVPGSAT-----CIVIKPMIYNFIQVTSF--LPILPMTV 228
DB 189 TNKVSFAVGVLMNRIPVSTDGIIVLHTEFCAMNOSRPDQKMIILFAFTVEFVIPAIA 248
QY 229 ISVLYYLMALRV 240
DB 249 IVIMYAHIAVOL 260

RESULT 15
093412 PRELIMINARY; PRT; 374 AA.
ID 093412;
AC 093412;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Orphan G protein-coupled receptor.
OS Spherooides nephelus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Spherooides.
OX NCBI_TaxID=39110;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20092336; PubMed=10628755;
RA Palyha O.C., Feigner S.D., Tan C.P., McKee K.K., Hreniuk D.L.,
RA Gao Y.D., Schlem K.D., Yang L., Morille G.J., Nargund R.,
RA Patchett A.A., Howard A.D., Smith R.G.;
RT "Ligand activation domain of human orphan growth hormone (GH)
RT secretagogue receptor (GHS-R) conserved from pufferfish to humans.";
RL Mol. Endocrinol. 14:160-169(2000).
DR EMBL; AF082209; AAC33472.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PRINTS; PR01565; NEUROMEDINR.
DR PRINTS; PR01566; NEUROMEDINDIR.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

```

```

DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
DR PROSITE; PS00043; HTH_GNTR_FAMILY; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 374 AA; 42324 MW; 2CF9304F004C7A16 CRC64;

Query Match 24.0%; Score 303; DB 13; Length 374;
Best Local Similarity 32.2%; Pred. No. 2.3e-20;
Matches 74; Conservative 51; Mismatches 71; Indels 34; Gaps 6;

QY 36 PRNSHFLP---VSVVYPIFVGVIGNVLCVLIQHQAARKTPTNYLFSLAASDLV 91
DB 26 PLNYYSIPLAVITVACTVILFTVGVGNVMTILVSRYRDMRTTNLYLCSMAVSDLF 85
QY 92 LILGMPLEVEEMRNYPFLFGPVGCYFKTALFETVCFASILSTTVSVERXVALHPFRA 151
DB 86 FVC-MPLDLYRMKRYRMRGDAICKLQFVSSESCYSTILCITALSVERRYLAICEPLRA 144
QY 152 KLOSTRRALRLIIGVWGFVSLPNTSINGIKFH--YPPNGSLVPGSA-----T 200
DB 145 KALVTKRVRALILLMTVSLNSGPFVGVAVYKDSIMFPNSDLNESSWPLEAVDTRE 204
QY 201 CIVIK-----PMIYNFIQVTSFLLPMTVISVLYYLMALRV 240
DB 205 CRMTQYAVESGLMEAMVWL-----SSVFEMPVCLTVLGLIGRRL 246

Search completed: January 17, 2003, 05:06:39
Job time : 97 secs

```



XX (PF1Z ) PFIZER LTD.  
 PA (PF1Z ) PFIZER INC.  
 XX  
 PI Harland L;  
 XX  
 DR WPI: 2001-302046/32;  
 DR P-PSDB: AAB68333.  
 XX  
 PT New human G-protein coupled receptor (GPCR) polynucleotides and  
 PT polypeptides, for screening modulators of the polypeptide useful in  
 PT treating diseases associated with signal transduction, e.g. cancer,  
 PT inflammation, or especially, obesity  
 PS  
 PS Claim 1: Page 42; 53pp; English.  
 XX  
 CC The present sequence encodes a human G-protein coupled receptor. The  
 CC G-protein coupled receptor polynucleotide and polypeptide are useful  
 CC as pharmaceuticals or in the manufacture of medicaments for the  
 CC treatment of obesity. They are useful in the diagnosis and treatment  
 CC of diseases and disorders associated with signal transduction such as  
 CC obesity, diabetes and metabolic disease, neurological disease, cancer,  
 CC psychotherapeutics, urogenital disease, inflammation, cancer, tissue  
 CC repair, dermatology, skin pigmentation, photoreaging, frailty,  
 CC osteoporosis, cardiovascular disease, gastrointestinal disease,  
 CC infection, allergy and respiratory disease, sensory organ disorders,  
 CC sleep disorders and hair loss. The polynucleotide may also be useful  
 CC in gene therapy.  
 XX  
 SQ Sequence 729 BP; 135 A; 245 C; 167 G; 182 T; 0 other:  
 Query Match 100.0%; Score 729; DB 22; Length 729;  
 Best Local Similarity 100.0%; Pred. No. 4e-148;  
 Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 541 ATCCATGGCATCAAGTTCCACTACTTCCCATGGTCCCTGGTCCAGGTGGCCACC 600  
 QY 601 TGTACGGTATCAAGCCCATGTGGATCTACAAATTTATCATCCAGTCACTCTTCCTA 660  
 Db 601 TGTACGGTATCAAGCCCATGTGGATCTACAAATTTATCATCCAGTCACTCTTCCTA 660  
 QY 661 TTCTACCTCCCTCCCAAGACGTGTCATCACTGTCCTACTACCTCATGGACACAGAGT 720  
 Db 661 TTCTACCTCCCTCCCAAGACGTGTCATCACTGTCCTACTACTACTCATGGACACAGAGT 720  
 QY 721 AGTATCTAG 729  
 Db 721 AGTATCTAG 729  
 RESULT 2  
 AAH50977/c  
 ID AAH50977 standard; DNA; 801 BP.  
 XX  
 AC AAH50977;  
 XX  
 DT 28-AUG-2001 (first entry)  
 XX  
 DE Human nGPCR15 coding sequence.  
 XX  
 KW G-protein-coupled receptor; nGPCR; seven transmembrane receptor;  
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;  
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;  
 KW cardiovascular disease; proliferative disorder; hormonal disorder;  
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;  
 KW attention deficit-hyperactivity disorder/attention deficit disorder;  
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;  
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;  
 KW neuroprotective; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200136473-A2.  
 XX  
 PD 25-MAY-2001.  
 XX  
 PF 16-NOV-2000; 2000WO-US31581.  
 XX  
 PR 16-NOV-1999; 99US-0165838.  
 PR 17-NOV-1999; 99US-0166071.  
 PR 19-NOV-1999; 99US-0166678.  
 PR 28-DEC-1999; 99US-0173396.  
 PR 22-FEB-2000; 2000US-0184129.  
 PR 28-FEB-2000; 2000US-0185421.  
 PR 28-FEB-2000; 2000US-0185534.  
 PR 02-MAR-2000; 2000US-0186530.  
 PR 03-MAR-2000; 2000US-0186811.  
 PR 09-MAR-2000; 2000US-0188114.  
 PR 17-MAR-2000; 2000US-0190310.  
 PR 21-MAR-2000; 2000US-0190800.  
 PR 20-APR-2000; 2000US-0198568.  
 PR 02-MAY-2000; 2000US-0201190.  
 PR 08-MAY-2000; 2000US-0203111.  
 PR 25-MAY-2000; 2000US-0207094.  
 XX  
 PA (PHAA ) PHARMACIA & UPJOHN CO.  
 PI  
 PI Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;  
 PI Schellin KA, Kayles PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;  
 DR WPI: 2001-389826/41.  
 DR P-PSDB: AAG80937.  
 XX  
 PT New G-protein-coupled receptor (nGPCR-x) and its encoding  
 PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -  
 XX  
 PS Claim 4: Page 78-79; 261pp; English.

CC The present invention relates to novel G protein-coupled receptors  
 CC (ngpRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,  
 CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present  
 CC sequence is the coding sequence for one such G protein-coupled receptor.  
 CC GPCRs are also known as seven transmembrane receptors and function in  
 CC signal transduction. The ngpRx coding sequences are useful for  
 CC screening a human to diagnose a disorder affecting the brain or a genetic  
 CC predisposition, specifically schizophrenia. ngpRx are useful for  
 CC identifying compounds useful for treating schizophrenia. Detection of  
 CC ngpRx in a sample is useful as a diagnostic tool for diseases or  
 CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,  
 CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular  
 CC diseases, proliferative disorders and hormonal disorders. Modulators of  
 CC ngpRx activity have the utility for treating neurological disorders,  
 CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity  
 CC disorder/attention deficit disorder), and neuronal disorders such as  
 CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.  
 CC Additional disorders include inflammatory conditions (e.g. Crohn's  
 CC disease), rheumatoid arthritis, autoimmune disorders, cancers,  
 CC respiratory ailments such as asthma, and inflammatory diseases e.g.  
 CC inflammatory bowel disease.

SO Sequence 801 BP; 200 A; 187 C; 261 G; 153 T; 0 other.

Query Match 100.0%; Score 729; DB 22; Length 801;  
 Best Local Similarity 100.0%; Pred. No. 4,1e-148;

Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAAACTCAGAAATGCTTCCTGGATCTACACAGAGAACTAGAAATCCATTCAG 60  
 DB 795 ATGGAAAACTCAGAAATGCTTCCTGGATCTACACAGAGAACTAGAAATCCATTCAG 736  
 QY 61 AAACACCTGAACAGCAGCCGAGGATATCTGSCCTTCCTCCGCGACCTCGCCAGCAC 120  
 DB 735 AAACACCTGAACAGCAGCCGAGGATATCTGSCCTTCCTCCGCGACCTCGCCAGCAC 676  
 QY 121 TTCTCTCCCGCGTGTCTGTGTGTATGTGCCAATTTTGTGGTGGGGCTCATTGGCAAT 180  
 DB 675 TTCTCTCCCGCGTGTGTGTGTGTATGTGCCAATTTTGTGGTGGGGCTCATTGGCAAT 616  
 QY 181 GTCTGTGTGTCTGGT 240  
 DB 615 GTCTGTGTGTCTGGT 556  
 QY 241 CTCTTCAGCCTGGCGGCTCTGACCTCTGACCTCTGCTCTCTTGGAAATGCCCTGGAGTTC 300  
 DB 555 CTCTTCAGCCTGGCGGCTCTGACCTCTGACCTCTGCTCTCTTGGAAATGCCCTGGAGTTC 496  
 QY 301 TATGAGATGTGGGCACTACCTTTCTTGTGGGCGCGTGGGCTGCTACTTCAAGACG 360  
 DB 495 TATGAGATGTGGGCACTACCTTTCTTGTGGGCGCGTGGGCTGCTACTTCAAGACG 436  
 QY 361 GCCCTCTTTAGACCGTGTGCTTGCCCTCATCCTCAGCATCAACACCGTACGGTGGAG 420  
 DB 435 GCCCTCTTTAGACCGTGTGCTTGCCCTCATCCTCAGCATCAACACCGTACGGTGGAG 376  
 QY 421 CGCTAGCTGGCAATCTACACCGCTTCCGGGCAAAATGAGAGACACCGCGCGCGGCC 480  
 DB 375 CGCTAGCTGGCAATCTACACCGCTTCCGGGCAAAATGAGAGACACCGCGCGCGGCC 316  
 QY 481 CTGAGATCTCTGGGCACTGTCTGGGCGTCTCGTGTCTTTCCTCCGCCAACACAGC 540  
 DB 315 CTGAGATCTCTGGGCACTGTCTGGGCGTCTCGTGTCTTTCCTCCGCCAACACAGC 256  
 QY 541 ATCCAGGCAATCAAGTTCCACTTCCCAATGGGCTCCGTGTCCAGGTTGGGCCACC 600  
 DB 255 ATCCAGGCAATCAAGTTCCACTTCCCAATGGGCTCCGTGTCCAGGTTGGGCCACC 136  
 QY 601 TGTAGGTCATCAAGCCCATGTGGATCTACATTTCTCATCTCCAGGCTACCTCTTCTTA 660  
 DB 195 TGTAGGTCATCAAGCCCATGTGGATCTACATTTCTCATCTCCAGGCTACCTCTTCTTA 136  
 QY 661 TTCTACCTCTCCCATGACTGTATCATGTGCTCTTACTACTCATGAGCTACAGAGTG 720

DB 135 TTCTACTCTCTCCCAAGACTGTATCATGATGTCTCTACTACTCATMGGACTACAGATG 76  
 QY 721 AGTATCTAG 729  
 DB 75 AGTATCTAG 67

#### RESULT 3

AA598102/c  
 ID AA598102 standard; DNA: 813 BP.

AA598102;

12-MAR-2002 (first entry)

Human DNA for potential G protein-coupled receptor #59.

Human: G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;  
 Alzheimer's disease; amyotrophic lateral sclerosis; asthma;  
 atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;  
 chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;  
 depression; epilepsy; macular degeneration; lymphoma; melanoma;  
 multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;  
 psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;  
 tuberculosis; cognition disorder; memory disorder; anorexia;  
 hormonal release disorder; cardiovascular activity disorder;  
 pain perception disorder; obesity; diabetes; diabetes; obesity;  
 diabetes; hyperlipidaemia; stroke; gene therapy.

Homo sapiens.

MO200185791-A1.

15-NOV-2001.

11-MAY-2001; 2001WO-US15332.

11-MAY-2000; 2000US-203217P.

18-MAY-2000; 2000US-205945P.

(LIFE-) LIFESPAN BIOSCIENCES INC.

Brown JP, Miller M, Burner G, Fabre-Suver C, Pritchard D;  
 WPI; 2002-066595/09.

Novel G protein-coupled receptor polypeptides including galanin  
 receptor polypeptides useful for identifying modulators that are useful  
 for treating Alzheimer's disease, psoriasis, melanoma, multiple  
 sclerosis, stroke -  
 Claim 2; Page 79; 144p; English.

The invention relates to an isolated polypeptide encoded by a  
 nucleic acid molecule that is at least 80% identical to the G  
 protein-coupled (GPCR) polynucleotides included in the specification.  
 Also included are probes based on the GPCR sequences (including  
 antisense probes), a host cell comprising an expression vector comprising  
 the GPCR sequence, antibodies raised against the polypeptides,  
 and methods of identifying modulators of the polypeptides. The  
 polypeptides are useful for identifying modulator compounds which  
 function as modulators, activators, repressors, agonists or antagonists  
 of the novel GPCR polypeptides including the GAL4 polypeptide. The  
 antibodies and nucleic acid probes as described above can be used to  
 detect the presence of the polypeptides and nucleic acids and are used to  
 diagnose a variety of diseases or disorders in which GPCRs are involved  
 e.g., Alzheimer's disease, amyotrophic lateral sclerosis, asthma,  
 atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy,  
 chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease,  
 depression, epilepsy, macular degeneration, lymphoma, melanoma,  
 multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease,  
 psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis,



CC tuberculosis and many other diseases listed in the specification. The  
 CC probes and antibodies are also useful for diagnosing cognition and memory  
 CC disorders, anorexia, hormonal release disorders, cardiovascular activity  
 CC disorders, pain perception disorders, obesity, diabetes, Alzheimer's  
 CC disease. Preferably, compounds that decrease or increase  
 CC the expression of galanin receptor (GAL4) can be used to treat obesity,  
 CC diabetes, hyperlipidaemia and stroke. The GPCR nucleic acid is  
 CC useful for treating the above mentioned disorders by gene therapy  
 CC techniques. The present sequence is a novel GPCR polynucleotide of the  
 CC invention.

XX Sequence 813 BP; 206 A; 190 C; 261 G; 156 T; 0 other;

Query Match 100.0%; Score 729; DB 24; Length 813;

Best Local Similarity 100.0%; Pred. No. 4,1e-148;

Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAAAACTTGAGAAATGCTTCCTGGATCTACACGAGAACTAGAAATCCATTCCAG 60

DB 795 ATGGAAAACTTGAGAAATGCTTCCTGGATCTACACGAGAACTAGAAATCCATTCCAG 736

QY 61 AACACCTGACAGCAGCAGGAGATCTGGCCCTCCTCGGAGCTCGGCGAGCCAC 120

DB 735 AACACCTGACAGCAGCAGGAGATCTGGCCCTCCTCGGAGCTCGGCGAGCCAC 676

QY 121 TTCTTCTCCCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 180

DB 675 TTCTTCTCCCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 616

QY 181 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240

DB 615 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556

QY 241 CTCTTTCAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

DB 555 CTCTTTCAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466

QY 301 TATGAGATGAGGAGCAATCACTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360

DB 495 TATGAGATGAGGAGCAATCACTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 456

QY 361 GCCCTCTTTAGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

DB 435 GCCCTCTTTAGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 376

QY 421 CGCTAGCTGCGACCTGACACCGCTGCGGCAAACTGAGAGCAACCGCGGCGCGCGCC 480

DB 375 CGCTAGCTGCGACCTGACACCGCTGCGGCAAACTGAGAGCAACCGCGGCGCGCGCC 316

QY 481 CTGAGATCTGCGGACCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

DB 315 CTGAGATCTGCGGACCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256

QY 541 ATCCATGAGCATCAAGTTCACATCTTCCCAATGAGGCTCCGTGCCAGTTCCGCGCAC 600

DB 255 ATCCATGAGCATCAAGTTCACATCTTCCCAATGAGGCTCCGTGCCAGTTCCGCGCAC 196

QY 601 TGTAGGATCAAGGCGCATGAGATCTACATTTCAATCTCAGAGCTACCTGCTTCTCTA 660

DB 195 TGTAGGATCAAGGCGCATGAGATCTACATTTCAATCTCAGAGCTACCTGCTTCTCTA 136

QY 661 TTCTACCTCTCCCAATGAGTGTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720

DB 135 TTCTACCTCTCCCAATGAGTGTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 76

QY 721 AGTATCTAG 729

DB 75 AGTATCTAG 67

RESULT 4  
 AAS98055  
 ID AAS98055 standard; DNA; 1239 BP.

XX  
 AC AAS98055;  
 XX  
 DT 12-MAR-2002 (first entry)  
 XX  
 XX Human DNA for potential G protein-coupled receptor #13.

Human: G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;  
 KW Alzheimer's disease; amyotrophic lateral sclerosis; asthma;  
 KW Chondrosarcoma; basal cell carcinoma; breast carcinoma; cardiomyopathy;  
 KW depression; epilepsy; macular degeneration; lymphoma; melanoma;  
 KW multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;  
 KW psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;  
 KW tuberculosis; cognition disorder; memory disorder; anorexia;  
 KW hormonal release disorder; cardiovascular activity disorder;  
 KW pain perception disorder; obesity; diabetes; obesity;  
 KW diabetes; hyperlipidaemia; stroke; gene therapy.

XX Homo sapiens.

XX MO200185791-A1.

XX 15-NOV-2001.

XX 11-MAY-2001; 2001MO-US15332.

XX 11-MAY-2000; 2000US-203217P.

XX 18-MAY-2000; 2000US-205945P.

XX (LIFE-) LIFESPAN BIOSCIENCES INC.

XX Brown JP, Miller M, Burner G, Fabre-Suver C, Pritchard D;  
 XX WPI; 2002-066595/09.

XX Novel G protein-coupled receptor polypeptides including galanin  
 PT receptor polypeptides useful for identifying modulators that are useful  
 PT for treating Alzheimer's disease, psoriasis, melanoma, multiple  
 PT sclerosis, stroke

XX Disclosure; Page 120-121; 144pp; English.

XX The invention relates to an isolated polypeptide encoded by a  
 CC nucleic acid molecule that is at least 80% identical to the G  
 CC protein-coupled (GPCR) polynucleotides included in the specification.

CC Also included are probes based on the GPCR sequences (including  
 CC antisense probes), a host cell comprising an expression vector comprising  
 CC the GPCR sequence, antibodies raised against the polypeptides,  
 CC and methods of identifying modulators of the polypeptides. The

CC polypeptides are useful for identifying modulator compounds which  
 CC function as modulators, activators, repressors, agonists or antagonists  
 CC of the novel GPCR polypeptides including the GAL4 polypeptide. The

CC antibodies and nucleic acid probes as described above can be used to  
 CC detect the presence of the polypeptides and nucleic acids and are used to  
 CC diagnose a variety of diseases or disorders in which GPCRs are involved

CC e.g. Alzheimer's disease, amyotrophic lateral sclerosis, asthma,  
 CC atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy,  
 CC chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease,  
 CC depression, epilepsy, macular degeneration, lymphoma, melanoma,

CC multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease,  
 CC psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis,  
 CC tuberculosis and many other diseases listed in the specification. The

CC probes and antibodies are also useful for diagnosing cognition and memory  
 CC disorders, anorexia, hormonal release disorders, cardiovascular activity  
 CC disorders, pain perception disorders, obesity, diabetes, Alzheimer's

CC disease. Preferably, compounds that decrease or increase  
 CC the expression of galanin receptor (GAL4) can be used to treat obesity,  
 CC diabetes, hyperlipidaemia and stroke. The GPCR nucleic acid is

CC useful for treating the above mentioned disorders by gene therapy  
 CC techniques. The present sequence is a novel GPCR polynucleotide of the  
 CC invention.



SQ Sequence 1239 BP; 267 A; 392 C; 265 G; 315 T; 0 other;

Query Match 98.5%; Score 717.8; DB 24; Length 1239;  
Best Local Similarity 99.7%; Pred. No. 1.2e-145;  
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAATTGAGATGCTTCCTGGATCTACAGAGAGAACTAGAGATTCATTCCAG 60  
DB 1 ATGAAAAAATTGAGATGCTTCCTGGATCTACAGAGAGAACTAGAGATTCATTCCAG 60  
QY 61 AAACACCTGAGACAGACCGAGAGATCTGGCCTTCTCTGGGAGCCTGGCGAGCCAC 120  
DB 61 AAACACCTGAGACAGACCGAGAGATCTGGCCTTCTCTGGGAGCCTGGCGAGCCAC 120  
121 TTCTTCCTCCCGTGTGTGTGTATGTGCAATTTTGTGGTGGGGGCTATTGGCAAT 180  
121 TTCTTCCTCCCGTGTGTGTGTATGTGCAATTTTGTGGTGGGGGCTATTGGCAAT 180  
QY 181 GTCCGTGTGCTGTGATCTGACAGACAGGCTATGAGAGCGCCCACTACTATAC 240  
DB 181 GTCCGTGTGCTGTGATCTGACAGACAGGCTATGAGAGCGCCCACTACTATAC 240  
QY 241 CTCTTCAGCCTGGGGGTCTCTGACCTCTGTCTCTCTCTGGAATGCCCTGGAGGTC 300  
DB 241 CTCTTCAGCCTGGGGGTCTCTGACCTCTGTCTCTCTCTGGAATGCCCTGGAGGTC 300  
QY 301 TATGAGATGTGGGCACTACCTCTTCTGTGGGCGCGCTGCTCTACTTCAAGACG 360  
DB 301 TATGAGATGTGGGCACTACCTCTTCTGTGGGCGCGCTGCTCTACTTCAAGACG 360  
QY 361 GCCCTTTTGGAGACCGTGTGCTTCCCTCCATCTCTAGATCAACACCTGAGGAG 420  
DB 361 GCCCTTTTGGAGACCGTGTGCTTCCCTCCATCTCTAGATCAACACCTGAGGAG 420  
QY 421 CGCTAGCTGGCCATCTCAACACCGGTTCCGGCCAACTGACAGACCGCGCGGCGCC 480  
DB 421 CGCTAGCTGGCCATCTCAACACCGGTTCCGGCCAACTGACAGACCGCGCGGCGCC 480  
QY 481 CTGAGATCTCTGGCATCTGTGGGGCTCTCGTGTCTTCTCCCTGGCCCAACAGC 540  
DB 481 CTGAGATCTCTGGCATCTGTGGGGCTCTCGTGTCTTCTCCCTGGCCCAACAGC 540  
QY 541 ATCCATGCGATCAAGTTCACACTTCCCAATGGGTCCCTGGTCCAGGTTGGCCACC 600  
DB 541 ATCCATGCGATCAAGTTCACACTTCCCAATGGGTCCCTGGTCCAGGTTGGCCACC 600  
QY 601 TGTACGTCATCAAGCCATGTGATCTACAAATTTCAATCCAGGTCACCTCTTCTA 660  
DB 601 TGTACGTCATCAAGCCATGTGATCTACAAATTTCAATCCAGGTCACCTCTTCTA 660  
QY 661 TTCTACCTCTCCCATGACTGTATCATGAGTGTCTCTACTACTGACGACGACGAGTG 720  
DB 661 TTCTACCTCTCCCATGACTGTATCATGAGTGTCTCTACTACTGACGACGACGAGTG 720  
QY 721 A 721  
DB 721 A 721

RESULT 5  
AAH43072  
ID AAH43072 standard; DNA: 1245 BP.  
XX  
AC AAH43072;  
XX  
DT 15-OCT-2001 (first entry)  
XX  
DE Nucleotide sequence of a human TGR-1 protein.  
XX  
KW TGR-1; neuromedin U; hypertension; stress disease; ss.  
OS Homo sapiens.  
XX

EH Key Location/Qualifiers  
FT CDS 1..1245  
FT /+lag= a  
FT /product= "TGR-1"  
FT /note= "no termination codon given"

MO200157524-A1.  
PD 09-AUG-2001.  
XX  
XX 02-FEB-2001; 2001MO-JP00746.  
XX  
XX 04-FEB-2000; 2000JP-0032773.  
XX  
XX 24-FEB-2000; 2000JP-0052252.  
XX  
XX 30-MAR-2000; 2000JP-0097896.  
XX  
XX 19-JUN-2000; 2000JP-0187536.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;  
XX Okubo S;  
XX  
XX WPI; 2001-488917/53.  
XX  
XX P-PSDB; AAC63353.  
XX  
XX  
XX PT Identifying predicted or actual structures of two or more members of a  
XX chemical or physical library by mass spectrometry compising  
XX correlating molecular mass measurements of two or more members with a  
XX shared chemical history -  
XX  
XX Claim 9; Page 79-80; 95pp; Japanese.

CC The present sequence encodes a human TGR-1 protein. The specification  
CC describes a method of screening a compound, which is capable of binding  
CC properties of neuromedin U to TGR-1. The method is useful for screening  
CC preventatives and remedies for hypertension, stress diseases, etc..  
CC TGR-1 antagonists are also useful for treating the same diseases.  
XX

SQ Sequence 1245 BP; 268 A; 393 C; 268 G; 316 T; 0 other;

Query Match 98.5%; Score 717.8; DB 22; Length 1245;  
Best Local Similarity 99.7%; Pred. No. 1.2e-145;  
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAATTGAGATGCTTCCTGGATCTACAGAGAGAACTAGAGATTCATTCCAG 60  
DB 10 ATGAAAAAATTGAGATGCTTCCTGGATCTACAGAGAGAACTAGAGATTCATTCCAG 69  
QY 61 AAACACCTGAGACAGACCGAGAGATCTGGCCTTCTCTGGGAGCCTGGCGAGCCAC 120  
DB 70 AAACACCTGAGACAGACCGAGAGATCTGGCCTTCTCTGGGAGCCTGGCGAGCCAC 129  
QY 121 TTCTTCCTCCCGTGTGTGTGTATGTGCAATTTTGTGGTGGGGGCTATTGGCAAT 180  
DB 130 TTCTTCCTCCCGTGTGTGTGTATGTGCAATTTTGTGGTGGGGGCTATTGGCAAT 189  
QY 181 GTCTGTGTGCTGTGATCTGACAGACAGGCTATGAGAGCGCCCACTACTATAC 240  
DB 190 GTCTGTGTGCTGTGATCTGACAGACAGGCTATGAGAGCGCCCACTACTATAC 249  
QY 241 CTCTTCAGCCTGGGGGTCTCTGACCTCTGTGCTCTCTTGGAAATGCCCTGGAGGTC 300  
DB 250 CTCTTCAGCCTGGGGGTCTCTGACCTCTGTGCTCTCTTGGAAATGCCCTGGAGGTC 309  
QY 301 TATGAGATGTGGGCACTACCTCTTCTGTGGGCGCGTGGGCTGCTCAATTCAGAGC 360  
DB 310 TATGAGATGTGGGCACTACCTCTTCTGTGGGCGCGTGGGCTGCTCAATTCAGAGC 369  
QY 361 GCCCTTTTGGAGACCGTGTGCTTCCCTCCATCTCTAGATCAACACCGTGGAGGAG 420  
DB 370 GCCCTTTTGGAGACCGTGTGCTTCCCTCCATCTCTAGATCAACACCGTGGAGGAG 429  
QY 421 CGCTAGCTGGCCATCTCAACACCGGTTCCGGGCCAACTGACAGACCGCGCGGCGCC 480

```

|||||
Db 430 CGCTAGCTGGCATTCTACACCCGTTCCGCGCAAACTGCAGACACCCGGCGGCC 489
QY 481 CTCAGAGTCTCGGCAATCGTGGGGCTCTCGTGTCTTCCCTGCCCAACACAGC 540
Db 490 CTCAGAGTCTCGGCAATCGTGGGGCTCTCGTGTCTTCCCTGCCCAACACAGC 549
QY 541 ATCCATGGCATCAAGTTCACACTTCCCAATGGGTCCTGGTCCAGTTGGCCACC 600
Db 550 ATCCATGGCATCAAGTTCACACTTCCCAATGGGTCCTGGTCCAGTTGGCCACC 609
QY 601 TGTACGGTCAATCAAGCCATGTGATCTACAAATTTATCAATCCAGGTACCTCTTCTTA 660
Db 610 TGTACGGTCAATCAAGCCATGTGATCTACAAATTTATCAATCCAGGTACCTCTTCTTA 669
QY 661 TTCTACCTCTCCCAATGATGTATCAATGATGTCTACATCCATGGCAGTCAAGAGT 720
Db 670 TTCTACCTCTCCCAATGATGTATCAATGATGTCTACATCCATGGCAGTCAAGACTA 729
QY 721 A 721
Db 730 A 730

RESULT 6
AAH43075 standard; DNA; 1245 BP.
AAH43075:
15-OCT-2001 (first entry)
Nucleotide sequence of a human TGR-1 protein.
TGR-1; neuromedin U; hypertension; stress disease; ss.
Homo sapiens.
Key Location/Qualifiers
CDS 1..1245
FT /tag= a
FT /product= "TGR-1"
FT /note= "no termination codon given"
MO200157524-A1.
09-AUG-2001.
02-FEB-2001; 2001WO-JP00746.
04-FEB-2000; 2000JP-0032773.
24-FEB-2000; 2000JP-0052252.
30-MAR-2000; 2000JP-0097896.
19-JUN-2000; 2000JP-0187536.
(TAKE ) TAKEDA CHEM IND LTD.
Hiruma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H,
Okubo S,
WPI: 2001-488917/53.
P-PSDB; AAG63366.
Identifying predicted or actual structures of two or more members of a
chemical or physical library by mass spectrometry comprising
correlating molecular mass measurements of two or more members with a
shared chemical history -
Disclosure; Page 85-87; 95pp; Japanese.
The present sequence encodes a human TGR-1 protein. The specification
describes a method of screening a compound, which is capable of binding
properties of neuromedin U to TGR-1. The method is useful for screening

```

```

CC Preventatives and remedies for hypertension, stress diseases, etc..
CC TGR-1 antagonists are also useful for treating the same diseases.
xx
SQ Sequence 1245 BP; 268 A; 394 C; 268 G; 315 T; 0 other;
Query Match 98.5%; Score 717.8; DB 22; Length 1245;
Best Local Similarity 99.7%; Pred. No. 1.2e-145;
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGAATAAATCTAGAAATGCTTCTCTGATCTACACAGAAATTAAGAATCTTCAG 60
Db 10 ATGGAATAAATCTAGAAATGCTTCTCTGATCTACACAGAAATTAAGAATCTTCAG 69
QY 61 AAACACCTGAACAGCAGCCGAGAGTATGAGCCCTTCTGCGAGCTCGGCGCAGCCAC 120
Db 70 AAACACCTGAACAGCAGCCGAGAGTATGAGCCCTTCTGCGAGCTCGGCGCAGCCAC 129
QY 121 TTCTTCCCTCCCGTGTCTGTGTATGATGATGATGATGATGATGATGATGATGAT 180
Db 130 TTCTTCCCTCCCGTGTCTGTGTATGATGATGATGATGATGATGATGATGATGAT 189
QY 181 GTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db 190 GTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 249
QY 241 CTCTTACGCTGGGCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Db 250 CTCTTACGCTGGGCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 309
QY 301 TATGAGATGTGGGCAACATCACTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Db 310 TATGAGATGTGGGCAACATCACTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 369
QY 361 GCCCTTTTGAAGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
Db 370 GCCCTTTTGAAGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 429
QY 421 CGCTAGCTGGCATTCTACACCCGTTCCGCGCAAACTGCAGACACCCGGCGGCC 480
Db 430 CGCTAGCTGGCATTCTACACCCGTTCCGCGCAAACTGCAGACACCCGGCGGCC 489
QY 481 CTCAGAGTCTCGGCAATCGTGGGGCTCTCGTGTCTTCCCTGCCCAACACAGC 540
Db 490 CTCAGAGTCTCGGCAATCGTGGGGCTCTCGTGTCTTCCCTGCCCAACACAGC 549
QY 541 ATCCATGGCATCAAGTTCACACTTCCCAATGGGTCCTGGTCCAGTTGGCCACC 600
Db 550 ATCCATGGCATCAAGTTCACACTTCCCAATGGGTCCTGGTCCAGTTGGCCACC 609
QY 601 TGTACGGTCAATCAAGCCATGTGATCTACAAATTTATCAATCCAGGTACCTCTTCTTA 660
Db 610 TGTACGGTCAATCAAGCCATGTGATCTACAAATTTATCAATCCAGGTACCTCTTCTTA 669
QY 661 TTCTACCTCTCCCAATGATGTATCAATGATGTCTACATCCATGGCAGTCAAGAGT 720
Db 670 TTCTACCTCTCCCAATGATGTATCAATGATGTCTACATCCATGGCAGTCAAGACTA 729
QY 721 A 721
Db 730 A 730

RESULT 7
AAD01123
ID AAD01123 standard; cDNA; 1248 BP.
AAD01123;
02-NOV-2000 (first entry)
Human orphan G protein-coupled receptor hRUP6 cDNA.
Human; orphan G protein-coupled receptor; GPCR; hRUP6; drug screening;
KW

```

XX	transmembrane receptor; signal cascade; ss.	
OS	Homo sapiens.	
XX		
XX	Key	Location/Qualifiers
PH	CDS	1..1248
FT		/*tag= a
FT		/product= "hrup6"
FT		/note= "Human orphan G protein-coupled receptor"
XX		
PN	MO200031258-A2.	
XX		
PD	02-JUN-2000.	
XX		
XX	13-OCT-1999;	99WO-US23687.
XX		
XX	20-NOV-1998;	98US-0109213.
XX	16-FEB-1999;	99US-0120416.
PR	26-FEB-1999;	99US-0121852.
PR	12-MAR-1989;	99US-0123946.
PR	12-MAR-1999;	99US-0123949.
PR	28-MAY-1999;	99US-0136436.
PR	28-MAY-1999;	99US-0136437.
PR	28-MAY-1999;	99US-0136439.
PR	28-MAY-1999;	99US-0136567.
PR	28-MAY-1999;	99US-0137127.
PR	28-MAY-1999;	99US-0137131.
PR	29-JUN-1999;	99US-0141448.
PR	29-SEP-1999;	99US-0156555.
PR	29-SEP-1999;	99US-0156633.
PR	29-SEP-1999;	99US-0156634.
PR	29-SEP-1999;	99US-0156653.
PR	01-OCT-1989;	99US-0157280.
PR	01-OCT-1989;	99US-0157281.
PR	01-OCT-1999;	99US-0157282.
PR	01-OCT-1999;	99US-0157293.
PR	01-OCT-1999;	99US-0157294.
PR	12-OCT-1999;	99US-0416760.
PR	12-OCT-1999;	99US-0417044.
XX		
PA	(AREN-) ARENA PHARM INC.	
XX		
XX	Chen R, Dang HT, Liaw CW, Lin I;	
XX	WPI: 2000-400068/34.	
DR	P-PSDB: AAY71296.	
XX		
XX	Novel human orphan G protein-coupled receptors and the encoding cDNAs for use in the identification of G protein-coupled receptor agonists -	
XX	Claim 21; Page 56-57; 102pp; English.	
XX		
XX	The present sequence is a cDNA encoding hrup6, an endogenous human	
XX	orphan G protein-coupled receptor (GPCR). The full length hrup6 cDNA	
CC	was cloned by RT-PCR using human thymus cDNA as template.	
CC	The orphan GPCR of the invention, like all GPCRs has seven transmembrane	
CC	alpha helices with an extracellular N-terminus and an intracellular	
CC	C-terminus. However, no endogenous ligands has yet been identified for	
CC	the proteins of the invention. The orphan GPCRs may be used in the	
CC	identification of their endogenous ligands, and to screen potential GPCR	
CC	agonists and antagonists for use as pharmaceutical agents. The proteins	
CC	may also be used in the study of GPCR-mediated signalling cascades, and	
CC	to elucidate their precise role in normal and diseased human conditions.	
CC	Nucleic acid encoding human orphan GPCRs may be used for tissue	
CC	localisation expression analysis to provide information about their	
CC	function in healthy and pathological states.	
XX		
XX		
XX	Sequence 1248 BP: 269 A; 393 C; 269 G; 317 T; 0 other;	
XX		
XX	Query Match	98.5%; Score 717.8; DB 21; Length 1248;
XX	Best Local Similarity	99.7%; Pred. No. 1.2e-145;
XX	Matches 719; Conservative	0; Mismatches 2; Indels 0; Gaps 0

[illegible]



XX Claim 1: Page 105; 137bp; Japanese.

PS The present sequence is the coding sequence for a human guanosine  
 CC triphosphate (GTP)-binding protein-coupled receptor. The receptor is  
 CC useful for the investigation, diagnosis, treatment and prevention of  
 CC diseases associated with GTP-binding protein-coupled receptors, including  
 CC neurological, circulatory, digestive system, immune system, muscle and  
 CC urinary system disorders. GTP-binding proteins are also known as  
 CC G-proteins.

XX Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other:

Query Match 98.5%; Score 717.8; DB 22; Length 1248;

Best Local Similarity 99.7%; Pred. No. 1.2e-145; Mismatches 2; Indels 0; Gaps 0;

atches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAAAACTTCAGATCTTCTGGATCTACACAGAACTAGAGATCCATTCCAG 60  
 DB 10 ATGGAAAACTTCAGATCTTCTGGATCTACACAGAACTAGAGATCCATTCCAG 69

QY 61 AAACACCTGAACGACGACGAGGATCTGGCTTCTCGGAGCTCGGCGACGCCAC 120  
 DB 70 AAACACCTGAACGACGACGAGGATCTGGCTTCTCGGAGCTCGGCGACGCCAC 129

QY 121 TTCTTCTCCCGCTCTGTGTGATGTGCAATTTTGTGGTGGGGTCATTGGCAAT 180  
 DB 130 TTCTTCTCCCGCTCTGTGTGATGTGCAATTTTGTGGTGGGGTCATTGGCAAT 189

QY 181 GTCTGTGGTGTCTGTGTATTTCTGACACACGAGGTATGAGAGCCCAACTACTAC 240  
 DB 190 GTCTGTGGTGTCTGTGTATTTCTGACACACGAGGTATGAGAGCCCAACTACTAC 249

QY 241 CTCTTACGCTGGGCGTCTGTGACCTCGTGTCTCTCTCTGATGTCCTTGGAGGTC 300  
 DB 250 CTCTTACGCTGGGCGTCTGTGACCTCGTGTCTCTCTCTGATGTCCTTGGAGGTC 309

QY 301 TATGAGATGTGGGCACTACCTTTCTGTGGGCGCGTGGGCTGACTTCAAGACG 360  
 DB 310 TATGAGATGTGGGCACTACCTTTCTGTGGGCGCGTGGGCTGACTTCAAGACG 369

QY 361 GCCCTCTTGAGACCGTGTGCTTGGCTCCATCCTCAGATCAGACCGTACGCGT 420  
 DB 370 GCCCTCTTGAGACCGTGTGCTTGGCTCCATCCTCAGATCAGACCGTACGCGT 429

QY 421 CGGTAGTGGCATCTTAACCCGTTCCGGCGCAAACTGACAGACCGCGCGCGGCC 480  
 DB 430 CGGTAGTGGCATCTTAACCCGTTCCGGCGCAAACTGACAGACCGCGCGCGGCC 489

QY 481 CTCAGATCTCGGCGATGCTGTGGGCTTCTCGTCTCTTCCCTGGCCCAACACG 540  
 DB 490 CTCAGATCTCGGCGATGCTGTGGGCTTCTCGTCTCTTCCCTGGCCCAACACG 549

QY 541 ATTCATGAGATCAAGTTCCTACTTCCCAATGGGTCCCTGTGCCAGTTTGGCGACC 600  
 DB 550 ATTCATGAGATCAAGTTCCTACTTCCCAATGGGTCCCTGTGCCAGTTTGGCGACC 609

QY 601 TGTAGGTCATCAAGCCCATGTGAGTCTACAAATTTATCATTCAGGTACCTCTTCT 660  
 DB 610 TGTAGGTCATCAAGCCCATGTGAGTCTACAAATTTATCATTCAGGTACCTCTTCT 669

QY 661 TTCTTACCTCTCCCATGAGTGTATGATGCTTACTTACCTCAATGGAGTCAAGAGT 720  
 DB 670 TTCTTACCTCTCCCATGAGTGTATGATGCTTACTTACTCTCATGTGACACTCAGACT 729

QY 721 A 721  
 DB 730 A 730

RESULT 10  
 ABK47909  
 ID ABK47909 standard; cDNA: 1248 BP.

XX ABK47909;  
 AC 18-JUN-2002 (first entry)  
 XX  
 DT  
 XX  
 DE Human G-protein coupled receptor AXOR34 cDNA.

KW Human; G protein-coupled; AXOR34; gene; ss; bacterial infection; cancer;  
 KW fungal infection; protozoan infection; viral infection; hypertension;  
 KW pain; diabetes; obesity; anorexia; bulimia; Parkinson's disease; stroke;  
 KW acute heart failure; hypotension; urinary retention; osteoporosis; ulcer;  
 KW angina pectoris; myocardial infarction; benign prostatic hypertrophy;  
 KW allergy; migraine; vomiting; psychotic disorder; neurological disorder;  
 KW dyskinesia; receptor.

XX Homo sapiens.

OS Key Location/Qualifiers

FT CDS 1..1248

FT /\*tag= a

FT /\*product= "Human AXOR34"

PN GB2363793-A.

PD 09-JAN-2002.

PF 10-APR-2001: 2001GB-0008983.

PR 10-APR-2000: 2000US-0545944.

PA (SMK ) SMITHKLINE BEECHAM CORP.

PI (SMK ) SMITHKLINE BEECHAM PLC.

PI Ames RS, Sarau H, Vawter L, Elshourbagy N, Shabon U;  
 PI Michalovich D;

DR WPI: 2002-306791/35.

DR P-PSDB; AAU77155.

PT New G protein-coupled receptor AXOR34, useful for diagnosis and

PT treatment of e.g. infections and cancer

XX Claim 1; Fig 1: 46pp; English.

XX The invention relates to a human G protein-coupled receptor AXOR34 and  
 CC the polynucleotide encoding it. The sequences of the invention are used  
 CC for treating and/or a variety of diseases, including bacterial, fungal,  
 CC protozoan and viral infections (particularly infections caused by HIV),  
 CC pain, cancers, diabetes, obesity, anorexia, bulimia, Parkinson's disease,  
 CC acute heart failure, hypotension, hypertension, urinary retention,  
 CC osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers,  
 CC allergies, benign prostatic hypertrophy, migraine, vomiting, psychotic  
 CC and neurological disorders (including anxiety, schizophrenia, manic  
 CC depression, depression, delirium, dementia and severe mental retardation)  
 CC and dyskinesias (such as Huntington's disease and Gilles de la Tourette's  
 CC syndrome). This sequence represents cDNA encoding the human AXOR34  
 CC polypeptide.

XX Sequence 1248 BP; 269 A; 392 C; 269 G; 318 T; 0 other:

Query Match 98.5%; Score 717.8; DB 24; Length 1248;

Best Local Similarity 99.7%; Pred. No. 1.2e-145;

Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAAAACTTCAGATCTTCTGGATCTACACAGAACTAGAGATCCATTCCAG 60  
 DB 10 ATGGAAAACTTCAGATCTTCTGGATCTACACAGAACTAGAGATCCATTCCAG 69

QY 61 AAACACCTGAACGACGACGAGGATCTGGCTTCTCGGAGCTCGGCGACGCCAC 120  
 DB 70 AAACACCTGAACGACGACGAGGATCTGGCTTCTCGGAGCTCGGCGACGCCAC 129

QY 121 TTCTTCTCCCGCTCTGTGTGATGTGCAATTTTGTGGTGGGGTCATTGGCAAT 180

```

Db      130 TTCTTCCCTCCCGCTGCTGTGATGTAGCCAAATTTTGTGTGGGGGTCATTGGCAAT 189
OY      181 GTCCCTGGTGGCTGGTGTATTCGACGACACAGGCTATGAAGCGCCCACTACTAC 240
Db      190 GTCCCTGGTGGCTGGTGTATTCGACGACACAGGCTATGAAGCGCCCACTACTAC 249
OY      241 CTCTTACGCTGGCGGCTCTGTGACCTCGAGGTCCTCTCTTGGATGGCCCGGAGGTC 300
Db      250 CTCTTACGCTGGCGGCTCTGTGACCTCGAGGTCCTCTCTTGGATGGCCCGGAGGTC 309
OY      301 TATGAGATGTGGCGCACTACCTTTCTTGTTCGGGCGCGTGGGCTACTTCAAGACG 360
Db      310 TATGAGATGTGGCGCACTACCTTTCTTGTTCGGGCGCGTGGGCTACTTCAAGACG 369
OY      361 GGCCTTTTGGAGCGGTGGCTGGCTGCATCCATCCATGACATCCACCGCTACCGTGGAG 420
Db      370 GGCCTTTTGGAGCGGTGGCTGGCTGCATCCATCCATGACATCCACCGCTACCGTGGAG 429
OY      421 CGCTACGTGGCATCTCTACACCGCTTCGCGCAACTGACAGACCGCGCGCGGCGC 480
Db      430 CGCTACGTGGCATCTCTACACCGCTTCGCGCAACTGACAGACCGCGCGCGGCGC 489
OY      481 CTCAGGATCTCTGGGCAATGCTGTGGGGCTTCGCTCTCTCTCTGCCCCAACACACAG 540
Db      490 CTCAGGATCTCTGGGCAATGCTGTGGGGCTTCGCTCTCTCTCTGCCCCAACACACAG 549
OY      541 ATCCAGGATCAAGTTCACACTCTTCCCCAATGGGTCCCTGGTCCAGTTCGGCGACC 600
Db      550 ATCCAGGATCAAGTTCACACTCTTCCCCAATGGGTCCCTGGTCCAGTTCGGCGACC 609
OY      601 TGTAGGTCATCAAGCCCATGTGATCTACAAATTTTCATCATCCAGGTCTCACTCTTA 660
Db      610 TGTAGGTCATCAAGCCCATGTGATCTACAAATTTTCATCATCCAGGTCTCACTCTTA 669
OY      661 TTCTACCTCTCCCATGACTGTGATCTACTACTCTATGCGACTCAAGATG 720
Db      670 TTCTACCTCTCCCATGACTGTGATCTACTACTCTATGCGACTCAAGATG 729
OY      721 A 721
Db      730 A 730

RESULT 11
AAD08008
ID      AAD08008 standard: cDNA: 1298 BP.
XX
AC      AAD08008;
XX
DT      07-AUG-2001 (first entry)
XX
DE      Human G-protein coupled receptor, SNORF72 cDNA.

Human; G-protein coupled receptor; SNORF72; neuromedin U neuropeptide;
NMU; inflammation; arthritis; autoimmune disease; septicemia; psychotic;
mental retardation; transplant rejection; neurological disorder; anxiety; pain;
respiratory disorder; depression; schizophrenia; dementia; obesity; pain;
gastrointestinal disorder; hypertension; hypotension; epilepsy; diabetes;
ischaemia; stroke; cancer; sexual disorder; circadian disorder; anorexia;
dermatological; psoriasis; Parkinson's disease; nausea; bulimia; allergy;
Alzheimer's disease; AIDS; hormonal disorder; memory disorder; migraine;
cardiovascular disorder; renal disorder; bone disease; delirium; asthma;
Cushing's disease; dysmenorrhoea; antiangiinal; cytostatic; osteoporosis;
metabolic disorder; behavioural disorder; Addison's disease; dyskinesia;
traquilliser; antidiabetic; antiaddictive; ss.

```

```

XX      XX
PN      MO200144297-A1.
XX      21-JUN-2001.
PD      13-DEC-2000; 2000MO-US33787.
XX      17-DEC-1999; 99US-0466435.
PR      25-APR-2000; 2000US-0558099.
PR      30-JUN-2000; 2000US-0609146.
XX      (SYNA-) SYNAPTRIC PHARM CORP.
PA      Bonini JA, Lerman GS, Quan Y, Ogozalek K;
PI      WPI: 2001-390240/41.
XX      P-PSDB: AAE03629.
DR      A purified mammalian SNORF62 or SNORF72 receptor protein for
PT      identification of compounds to treat e.g. inflammation, arthritis,
PT      autoimmune diseases, transplant rejection, AIDS, cancer -
XX      Claim 42; Fig 3; 256pp; English.
PS      The invention relates to human G-protein coupled receptors, SNORF62 and
CC      SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72
CC      receptors are specific for neuromedin U (NMU) neuropeptides, hence they
CC      are also known as NMU receptors. The agonist and antagonist of NMU
CC      receptors are useful for treating an abnormality in a subject that is
CC      alleviated by decreasing or increasing the activity of NMU receptor.
CC      The NMU receptors serves as a valuable tool for designing drugs which are
CC      useful for treating various pathophysiological conditions such as
CC      inflammation, arthritis, autoimmune diseases, transplant rejection,
CC      graft vs host disease, bacterial, fungal, protozoan and viral infections,
CC      septicemia, AIDS, pain, psychotic and neurological disorders, including
CC      anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC      loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
CC      eating/body weight disorders including obesity, bulimia, diabetes,
CC      anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
CC      disorders, ischaemia, stroke, cancer, sexual disorders, circadian
CC      disorders, renal disorders, bone diseases including osteoporosis, benign
CC      prostatic hypertrophy, gastrointestinal disorders, nasal congestion,
CC      dermatological disorders such as psoriasis, allergies, Parkinson's
CC      disease, Alzheimer's disease, acute heart failure, angina disorders,
CC      delirium and dyskinesias such as Huntington's disease. They can also be
CC      used to regulate steroid hormone disorders, epinephrine release
CC      disorders, electrolyte balance disorders, endocrine disorders, memory
CC      disorders, somatosensory disorders, metabolic disorders, behavioural
CC      disorders, drug addiction, migraine, Addison's disease, Cushing's
CC      disease, prevent miscarriage, induce labour or to treat dysmenorrhoea.
CC      The present cDNA sequence encodes human G-protein coupled receptor,
CC      SNORF72.
XX      Sequence 1298 BP: 278 A; 405 C; 282 G; 333 T; 0 other;
SQ

Query Match          98.5%; Score 717.8; DB 22; Length 1298;
Best Local Similarity 99.7%; Pred. No. 1,2e-145;
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 ATGGAAAACTTCAGAACTCTCTCTGATCTACACGACAAACTAGAGATCCATTCCAG 60
Db      36 ATGGAAAACTTCAGAACTCTCTCTGATCTACACGACAAACTAGAGATCCATTCCAG 95
OY      61 AAACACCTTAACAGACACGACGAGATATCGGCTTCCTGCGGACCTGGCGGACGCCAC 120
Db      96 AAACACCTTAACAGACACGACGAGATATCGGCTTCCTGCGGACCTGGCGGACGCCAC 155
OY      121 TTCTTCTCTCCCGTGTGCTGTGATGTGCAATTTTGTGTGGGGGTCAATTGGCAAT 180
Db      156 TTCTTCTCTCCCGTGTGCTGTGATGTGCAATTTTGTGTGGGGGTCAATTGGCAAT 215
OY      181 GTCTTGTGTGCTGTGATGTGACAGACACAGGCTATGAAGCGCCCACTACTACTAC 240

```

```

Db 216 GTCTGGTGTGCTGTGATTTGTGACAGCACGAGGTATGAGAGCCCAACTACTAC 275
QY 241 CTCTTGGAGCTGGAGGCTCTGTGACCTCTGATGCTCTCTTGGATGCCCCGAGAGTC 300
Db 276 CTCTTGGAGCTGGAGGCTCTGTGACCTCTGATGCTCTCTTGGATGCCCCGAGAGTC 335
QY 301 TATGAGATGTGGGCGCACTACCCCTTCTTGTGGGCGCGTGGAGTCTACTTCAAGAGC 360
Db 336 TATGAGATGTGGGCGCACTACCCCTTCTTGTGGGCGCGTGGAGTCTACTTCAAGAGC 395
QY 361 GCCCTCTTTGAGACCGTGTGCTTCCGCTCATCTCAGATCAGACCGTCAAGCTGGAG 420
Db 396 GCCCTCTTTGAGACCGTGTGCTTCCGCTCATCTCAGATCAGACCGTCAAGCTGGAG 455
421 CGCTAGTGGGCGCACTACCCGCTTCCGCGCAACTGAGAGCAGCCGGGCGCGGCGC 480
456 CGCTAGTGGGCGCACTACCCGCTTCCGCGCAACTGAGAGCAGCCGGGCGCGGCGC 515
QY 481 CTCAGATCTCGGCACTGCTGTGGGCGCTTCCGCTCTTCTCCCTGCCCAACACAGC 540
Db 516 CTCAGATCTCGGCACTGCTGTGGGCGCTTCCGCTCTTCTCCCTGCCCAACACAGC 575
QY 541 ATCCATGGATCAAGTTCACCTACTTCCCAATGGGTCCTGGTCCCAAGTTGGGCGAC 600
Db 576 ATCCATGGATCAAGTTCACCTACTTCCCAATGGGTCCTGGTCCCAAGTTGGGCGAC 635
QY 601 TGTACGGTCAAGCGCATGTGATCTACATTTCACTCCAGGTCACCTCTTCTCTA 660
Db 636 TGTACGGTCAAGCGCATGTGATCTACATTTCACTCCAGGTCACCTCTTCTCTA 695
QY 661 TTCTACCTCTCCCACTGACTGTCACTAGTGTCTACTACTACTGACATGACATGAGT 720
Db 696 TTCTACCTCTCCCACTGACTGTCACTAGTGTCTACTACTACTGACATGACATGAGT 755
QY 721 A 721
Db 756 A 756

```

RESULT 12  
AAD23701  
ID AAD23701 standard; cDNA; 1344 BP.

AAD23701:  
07-MAR-2002 (first entry)

Human NMUR2 cDNA.  
Human: neuromedin U receptor; NMUR2; FM-4 protein; eating; weight gain;  
feeding behaviour; ss.  
Homo sapiens.

Key Location/Qualifiers  
FT 55..1302  
FT /tag= a  
FT /product= "Human NMUR2 protein"

XX MO200181418-A1.  
XX  
XX  
XX 01-NOV-2001.  
XX  
XX 25-APR-2001: 2001MO-US13386.  
XX  
XX 27-APR-2000: 2000US-200718P.  
XX  
XX (MERL) MERCK & CO INC.  
XX  
XX Liu Q, Lynch KR, Howard AD, Mellin TN, Strack A;  
PI Van Der Ploeg LHT, Wang R, Jiang Q, Williams D;  
XX  
XX WPI; 2002-034435/04.  
XX  
XX

```

DR P-PSDB: AAE14262.
XX
XX New polypeptide, useful for identifying compounds which modulate the
PT feeding activity of a mammal, comprises the novel isolated human and
PT rat neuromedin U receptor designated NMUR2 -
XX
XX Example 1; Fig 1: 47pp; English.
XX
XX The invention relates to human and rat neuromedin U receptor designated
CC NMUR2 polypeptides and polynucleotides. NMUR2 also referred as FM-4, is
CC free from associated proteins and is involved in the feeding behaviour
CC in mammals. Sequences of the invention are useful for identifying
CC compounds which modulate the feeding activity of a mammal. The compounds
CC identified are useful for modulating eating and weight gain. The present
CC sequence is human NMUR2 cDNA.
XX
SQ Sequence 1344 BP; 291 A; 419 C; 290 G; 344 T; 0 other;

```

Query Match 98.5%; Score 717.8; DB 24; Length 1344;  
Best Local Similarity 99.7%; Pred. No. 1.2e-145;  
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGGAAGAACTTCAAGATGCTTCTCGATCTACACAGCAAACTAGAAATCATTCAC 60
Db 64 ATGGAAGAACTTCAAGATGCTTCTCGATCTACACAGCAAACTAGAAATCATTCAC 123
QY 61 AAGACCTGAGACGACCGGAGATATCGGCTTCTCGGAGACCTGGGCGGACGAC 120
Db 124 AAGACCTGAGACGACCGGAGATATCGGCTTCTCGGAGACCTGGGCGGACGAC 183
QY 121 TTCTTCCTCCCGCTGTCTGTGTATGTGCAATTTTGTGGTGGGGTCAATTGGCAAT 180
Db 184 TTCTTCCTCCCGCTGTCTGTGTATGTGCAATTTTGTGGTGGGGTCAATTGGCAAT 243
QY 181 GTCTGGTGTGCTGTGATTTGACAGCACAGGCTATGAGAGCCCAACTACTAC 240
Db 244 GTCTGGTGTGCTGTGATTTGACAGCACAGGCTATGAGAGCCCAACTACTAC 303
QY 241 CTCTTGGAGCTGGGCGTCTGTGACCTGTGCTCTCTCTTGGATGCCCCGAGGTC 300
Db 304 CTCTTGGAGCTGGGCGTCTGTGACCTGTGCTCTCTCTTGGATGCCCCGAGGTC 363
QY 301 TATGAGATGTGGGCGCACTACCTTTCTTGTGGGCGCGTGGGCTCTACTTCAAGAGC 360
Db 364 TATGAGATGTGGGCGCACTACCTTTCTTGTGGGCGCGTGGGCTCTACTTCAAGAGC 423
QY 361 GCCCTCTTTGAGACCGTGTGCTTCCGCTCATCTCAGATCAGACCGTCAAGCTGGAG 420
Db 424 GCCCTCTTTGAGACCGTGTGCTTCCGCTCATCTCAGATCAGACCGTCAAGCTGGAG 483
QY 421 CGCTAGTGGGCGCACTACCCGCTTCCGCGCAACTGAGAGCAGCCGGGCGCGGCGC 480
Db 484 CGCTAGTGGGCGCACTACCCGCTTCCGCGCAACTGAGAGCAGCCGGGCGCGGCGC 543
QY 481 CTCAGATCTCGGCACTGCTGTGGGCGCTTCCGCTCTTCTCCCTGCCCAACACAGC 540
Db 544 CTCAGATCTCGGCACTGCTGTGGGCGCTTCCGCTCTTCTCCCTGCCCAACACAGC 603
QY 541 ATCCATGGATCAAGTTCACCTACTTCCCAATGGGTCCTGGTCCAGGTTGGGCGAC 600
Db 604 ATCCATGGATCAAGTTCACCTACTTCCCAATGGGTCCTGGTCCAGGTTGGGCGAC 663
QY 601 TGTACGGTCAAGCGCATGTGATCTACATTTCACTCCAGGTCACCTCTTCTCTA 660
Db 664 TGTACGGTCAAGCGCATGTGATCTACATTTCACTCCAGGTCACCTCTTCTCTA 723
QY 661 TTCTACCTCTCCCACTGACTGTCACTAGTGTCTACTACTACTGACATGACATGAGT 720
Db 724 TTCTACCTCTCCCACTGACTGTCACTAGTGTCTACTACTACTGACATGACATGAGT 783
QY 721 A 721
Db 784 A 784

```





```

XX WPI: 2001-273568/28.
DR P-PSDB: AAB67806.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers
XX
PS Example 1b: Page 90-92; 102pp; English.
XX
CC The present sequence encodes a splice variant of the long version of a
CC human G-protein coupled receptor designated IGS4A. IGS4 exists in two
CC polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4
CC polynucleotides are useful for preventing, ameliorating or correcting
CC dysfunctions or diseases. These diseases include peripheral nervous
CC system, psychiatric and central nervous system disorders
CC (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's
CC disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases
CC (e.g. heart failure, angina pectoris, myocardial infarction or
CC hypertension), dyslipidemias, obesity, emesis, gastrointestinal
CC disorders (e.g. inflammatory bowel disease or motility disorders),
CC osteoporosis, inflammations, infections (e.g. bacterial, fungal,
CC protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 1594 BP; 370 A; 473 C; 332 G; 419 T; 0 other:
Query Match 98.5%; Score 717.8; DB 22; Length 1594;
Best Local Similarity 99.7%; Pred. No. 1,2e-145;
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 ATGGAAACTTGAAGATCTTCTGATCTACACAGAACTAGAGATCCATCCAG 60
64 ATGGAAACTTGAAGATCTTCTGATCTACACAGAACTAGAGATCCATCCAG 123
61 AAACACCTGAACAGACAGGAGATCTGACCTTCTGCGGACCTGCGGACGAC 120
124 AAACACCTGAACAGACAGGAGATCTGACCTTCTGCGGACCTGCGGACGAC 183
121 TTCTTCTCCCTGCTGCTGATGTCGCAATTTTGTGTGGGGGTCATGGCAAT 180
184 TTCTTCTCCCTGCTGCTGATGTCGCAATTTTGTGTGGGGGTCATGGCAAT 243
181 GTCTGTGCTGCTGCTGATGTCGCAAGCAGGCTATGAGAGGCCCAACTACTC 240
244 GTCTGTGCTGCTGCTGATGTCGCAAGCAGGCTATGAGAGGCCCAACTACTC 303
241 GTCTGTGCTGCTGCTGATGTCGCAAGCAGGCTATGAGAGGCCCAACTACTC 300
304 GTCTGTGCTGCTGCTGATGTCGCAAGCAGGCTATGAGAGGCCCAACTACTC 363
301 TATGAAATGTGGGCAACTACTTCTTGTGTGGGGGTCGCTGCTTCAAGAGC 360
364 TATGAAATGTGGGCAACTACTTCTTGTGTGGGGGTCGCTGCTTCAAGAGC 423
361 GGCCTTTTGAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
424 GGCCTTTTGAAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
421 CGGTACGTGGGCACTTACACCGCTTCCGGCCAAACTGAGAGCAACCGGCGCC 480
484 CGGTACGTGGGCACTTACACCGCTTCCGGCCAAACTGAGAGCAACCGGCGCC 543
481 CTCAGATCTCTGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
544 CTCAGATCTCTGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 603

```

```

OY 541 ATCATGATCAAGTTCACACTTCTCCCAATGGTCCCTGCTCCAGGTGGCCACC 600
DB 604 ATCATGATCAAGTTCACACTTCTCCCAATGGTCCCTGCTCCAGGTGGCCACC 663
OY 601 TGTAGGTCATCAAGCCATGTGATCTCAATTTTCATCATCCAGGTCACCTCTCCCA 660
DB 664 TGTAGGTCATCAAGCCATGTGATCTCAATTTTCATCATCCAGGTCACCTCTCCCA 723
OY 661 TTCTACCTCTCCCAATGTGATCTCAATTTTCATCATCCAGGTCACCTCTCCCA 720
DB 724 TTCTACCTCTCCCAATGTGATCTCAATTTTCATCATCCAGGTCACCTCTCCCA 783
OY 721 A 721
DB 784 A 784

RESULT 15
AAF80327
ID AAF80327 standard; DNA; 1594 BP.
AC AAF80327;
DE 29-JUN-2001 (first entry)
XX
XX Splice variant of G-protein coupled receptor IGS4A short version cDNA.
XX
XX Human: G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
XX nervous system disorder; psychiatric disorder; Parkinson's disease;
XX episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
XX cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
XX gastrointestinal disorder; myocardial infarction; hypertension; dyslipidemia;
XX inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
XX gynecological disorder; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 64..945
XX FT /tag= "a"
XX FT /product= "splice variant of G-protein coupled receptor
XX IGS4A short version"
XX PN MO200125269-A2.
XX
XX PD 12-APR-2001.
XX
XX PF 25-SEP-2000; 2000MO-EP09584.
XX
XX PR 24-SEP-1999; 99EP-0203140.
XX PR 24-SEP-1999; 99NL-1013140.
XX PR 28-JUL-2000; 2000EP-0202683.
XX PR 31-JUL-2000; 2000US-0222047.
XX
XX PA (SOLV ) SOLVAY PHARM BV.
XX
XX PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX WPI: 2001-273568/28.
XX DR P-PSDB; AAB67807.
XX
XX PT New G-protein coupled receptors and the polynucleotides encoding them,
XX useful for preventing, ameliorating or correcting nervous system
XX disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
XX or cancers
XX
XX Example 1b: Page 94-95; 102pp; English.
XX
XX The present sequence encodes a splice variant of the short version of a
XX human G-protein coupled receptor designated IGS4A. IGS4 exists in two
XX polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4

```

CC polynucleotides are useful for preventing, ameliorating or correcting  
CC dysfunctions or diseases. These diseases include peripheral nervous  
CC system, psychiatric and central nervous system disorders  
CC (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's  
CC disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases  
CC (e.g. heart failure, angina pectoris, myocardial infarction or  
CC hypertension), dyslipidemias, obesity, emesis, gastrointestinal  
CC disorders (e.g. inflammatory bowel disease or motility disorders),  
CC osteoporosis, inflammations, infections (e.g. bacterial, fungal,  
CC protozoan or viral), pain, cancers, immune disorders, allergies,  
CC sepsis or gynecological disorders. Agonists or antagonists of IG54 are  
CC effective with regard to disorders of the nervous system, including the  
CC central and peripheral nervous systems, disorders of the gastrointestinal  
CC system, cardiovascular system, skeletal muscle, thyroid, lung or  
CC genitourinary system, or immunological disease. The IG54 polynucleotides  
CC are useful as diagnostic reagents for detecting under expression,  
CC overexpression or altered expression of IG54.  
XX  
SQ Sequence 1594 BP; 370 A; 473 C; 332 G; 419 T; 0 other;

Query Match 98.5%; Score 717.8; DB 22; Length 1594;

Best Local Similarity 99.7%; Pred. No. 1.2e-145;

Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGGAAGAACTTGTGATGCTTCTGATCTACACAGAGAACTAGAGATCCATTCCAG 60  
64 ATGGAAGAACTTGTGATGCTTCTGATCTACACAGAGAACTAGAGATCCATTCCAG 123  
QY 61 AACACCTGTAACGACGACGAGAGATCTGCGCTTCTGCGGACCTGCGGACGAC 120  
DB 124 AACACCTGTAACGACGAGAGATCTGCGCTTCTGCGGACCTGCGGACGAC 183  
QY 121 TTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
DB 184 TTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243  
QY 181 GTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
DB 244 GTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303  
QY 241 CTGTTACGCTGCGGCTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
DB 304 CTGTTACGCTGCGGCTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363  
QY 301 TATGAGATGAGGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
DB 364 TATGAGATGAGGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423  
QY 361 GCCCTCTTGTAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 424 GCCCTCTTGTAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483  
QY 421 CGCTAGCTGCGCATCTCTACACCGCTTCCGCGCAAACTGCAAGACCGCGCGGCC 480  
DB 484 CGCTAGCTGCGCATCTCTACACCGCTTCCGCGCAAACTGCAAGACCGCGCGGCC 543  
QY 481 CTCAGATCTCTGCGCATCTCTGCGGCTTCTCGGCTTCTTCCCTGCCCAACACAGC 540  
DB 544 CTCAGATCTCTGCGCATCTCTGCGGCTTCTCGGCTTCTTCTTCCCTGCCCAACACAGC 603  
QY 541 ATCCATGGCATCAAGTCCACTCTTCCCAATGGGTCCCTGGTCCAGGTTCCGCCACC 600  
DB 604 ATCCATGGCATCAAGTCCACTCTTCCCAATGGGTCCCTGGTCCAGGTTCCGCCACC 663  
QY 601 TGTACGGTCAATCAAGCCATGATGATCAATTTATCATCCAGGTCACCTCTCTCTA 660  
DB 664 TGTACGGTCAATCAAGCCATGATGATCAATTTATCATCCAGGTCACCTCTCTCTA 723  
QY 661 TTCTACCTCTTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
DB 724 TTCTACCTCTTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 783  
QY 721 A 721

Db 784 A 784

Search completed: January 17, 2003, 02:34:05  
Job time : 272 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2003, 02:34:16 ; Search time 2748 Seconds

(without alignments)  
7720.504 Million cell updates/sec

Title: US-09-684-725-1

Perfect score: 729  
Sequence: 1 atggaataactcagatgc.....cactcagagtgtatctag 729

ring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_com:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pin:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_higo\_hum:\*  
40: em\_higo\_mus:\*  
41: em\_higo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	729	100.0	729	6	AX139107	AX139107 Sequence
2	729	100.0	801	6	AX147772	AX147772 Sequence
3	729	100.0	214267	9	AC008571	AC008571 Homo sapi
4	717.8	98.5	1239	9	AF242874	AF242874 Homo sapi
5	717.8	98.5	1248	9	AB041228	AB041228 Homo sapi
6	717.8	98.5	1298	9	AF272363	AF272363 Homo sapi
7	717.8	98.5	1594	6	AX109242	AX109242 Sequence
8	717.8	98.5	1594	6	AX109244	AX109244 Sequence
9	717.8	98.5	1658	6	AX109234	AX109234 Sequence
10	717.8	98.5	1658	6	AX109236	AX109236 Sequence
11	717.8	98.5	1658	6	AX109238	AX109238 Sequence
12	717.8	98.5	1658	6	AX109240	AX109240 Sequence
13	716.2	98.2	1248	9	AF292402	AF292402 Homo sapi
14	716.2	98.2	2069	9	BC016938	BC016938 Homo sapi
15	476.8	65.4	165392	2	AC122979	AC122979 Rattus no
16	476.8	65.4	178043	2	AC110442	AC110442 Rattus no
17	476.2	65.3	213942	2	AL662792	AL662792 Rattus no
18	471.4	64.7	1188	10	AF242875	AF242875 Rattus no
19	471.4	64.7	1188	10	AV057384	AV057384 Mus muscu
20	471.4	64.7	1314	10	AB041229	AB041229 Rattus no
21	282.4	38.7	3354	9	BC036543	BC036543 Homo sapi
22	280.8	38.5	828	9	HS06PCR1	AF044600 Homo sapi
23	280.8	38.5	1209	6	BD012783	BD012783 Screening
24	280.8	38.5	1209	23	BD008218	BD008218 Screening
25	280.8	38.5	1318	9	AF272362	AF272362 Homo sapi
26	280.8	38.5	168860	9	AC017104	AC017104 Homo sapi
27	277.6	38.1	2374	6	AX342663	AX342663 Sequence
28	255	35.0	1209	10	AF242873	AF242873 Rattus no
29	255	35.0	1239	10	AB038649	AB038649 Rattus no
30	255	35.0	171749	2	AC112440	AC112440 Rattus no
31	243	33.3	1218	10	AF044602	AF044602 Mus muscu
32	243	33.3	203177	2	AC102609	AC102609 Mus muscu
33	161.6	22.2	167102	2	AC118178	AC118178 Rattus no
34	159.6	21.9	75950	2	AC073449	AC073449 Homo sapi
35	134.8	18.5	1161	6	AX154591	AX154591 Sequence
36	134.8	18.5	1239	6	AX154589	AX154589 Sequence
37	134.8	18.5	2040	9	AF034632	AF034632 Homo sapi
38	134.8	18.5	163284	9	AL137000	AL137000 Human DNA
39	134.8	18.5	341560	2	AL596304	AL596304 Homo sapi
40	133.2	18.3	1676	5	AF082210	AF082210 Sphoeroid
41	131.2	18.0	692	9	AF339407	AF339407 Homo sapi
42	127	17.4	1254	6	E11480	E11480 cDNA encodl
43	127	17.4	1257	6	AX280911	AX280911 Sequence
44	127	17.4	4131	9	HSNEURA	X70070 H.sapiens m
45	125.8	17.3	615	9	HSA339459	AJ339459 Homo sapi

#### ALIGNMENTS

RESULT 1  
AX139107  
LOCUS AX139107 729 bp DNA  
DEFINITION Sequence 1 from Patent EP1090990.  
ACCESSION AX139107  
VERSION AX139107.1 GI:14274786  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 729)  
AUTHORS Harland, L.P.  
TITLE Human g-protein-coupled receptor  
JOURNAL Patent: EP 1090990-A 1 11-APR-2001;

FEATURES Pfizer Limited (GB) ; PFIZER INC. (US)  
Location/Qualifiers  
source 1. .729  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 135 a 245 c 167 g 182 t

ORIGIN

Query Match 100.0%; Score 729; DB 6; Length 729;  
Best Local Similarity 100.0%; Pred. No. 1.8e-127;  
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAATACTTGAATGCTTCTGATCTACACAGAACTAGAAGATCCATTCCAG 60  
|||||  
Db 1 ATGGAATACTTGAATGCTTCTGATCTACACAGAACTAGAAGATCCATTCCAG 60

QY 61 AACACCTGTAACAGACACCGAGATCTGCTTCTCTGCGGAGCTCGGCGACGAC 120  
|||||  
Db 61 AACACCTGTAACAGACACCGAGATCTGCTTCTCTGCGGAGCTCGGCGACGAC 120

QY 121 TTCTTCTCTCCCGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180  
|||||  
Db 121 TTCTTCTCTCCCGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180

QY 181 GTCTTGT 240  
|||||  
Db 181 GTCTTGT 240

QY 241 CTCTTGAAGCTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300  
|||||  
Db 241 CTCTTGAAGCTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300

QY 301 TATGAGATGTGGGCGCAACTACCTTTCTTGTGGGCGCGTGGGCTGCTACTTCAAGAG 360  
|||||  
Db 301 TATGAGATGTGGGCGCAACTACCTTTCTTGTGGGCGCGTGGGCTGCTACTTCAAGAG 360

QY 361 GCCCTCTTTGAGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420  
|||||  
Db 361 GCCCTCTTTGAGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420

QY 421 CGCTAGTGAGCGATCTGACACCGGCTTCCGCAACTGAGAGACCGGCGCGGCGGC 480  
|||||  
Db 421 CGCTAGTGAGCGATCTGACACCGGCTTCCGCAACTGAGAGACCGGCGCGGCGGC 480

QY 481 CTCAGGATCTCGGCAATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540  
|||||  
Db 481 CTCAGGATCTCGGCAATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540

QY 541 ATCCATGCGATCAAGTTCACACTTCCCAATGGGCTCCCTGTGCCAGGTTCCGCGACG 600  
|||||  
Db 541 ATCCATGCGATCAAGTTCACACTTCCCAATGGGCTCCCTGTGCCAGGTTCCGCGACG 600

QY 601 TGTAGGTCATCAAGCCCATGTGATCTCAATTTTATCATCTCCAGGTCACCTCTCTCTA 660  
|||||  
Db 601 TGTAGGTCATCAAGCCCATGTGATCTCAATTTTATCATCTCCAGGTCACCTCTCTCTA 660

QY 661 TTCTTACCTCTCCCGATGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 720  
|||||  
Db 661 TTCTTACCTCTCCCGATGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 720

QY 721 AGTATCTAG 729  
|||||  
Db 721 AGTATCTAG 729

RESULT 2  
AX147772/c 801 bp DNA linear PAT 08-JUN-2001  
LOCUS AX147772  
DEFINITION Sequence 17 from Patent W00136473.  
ACCESSION AX147772  
VERSION AX147772.1 GI:14346807  
KEYWORDS human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 801)  
Vogel, G., Wood, L.S., Parodi, L.A., Hibbsch, R.R., Lind, P.,  
Slightom, J., Schellin, K.A., Kayles, P.S., Bannigan, C.M., Ruff, V.,  
Sejltz, T., and Huff, R.M.  
Novel 9 protein-coupled receptors  
Patent: WO 0136473-A 17 25-MAY-2001;  
PHARMACIA & UPJOHN COMPANY (US)

FEATURES  
source 1. .801  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 200 a 187 c 261 g 153 t

ORIGIN

Query Match 100.0%; Score 729; DB 6; Length 801;  
Best Local Similarity 100.0%; Pred. No. 1.8e-127;  
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAATACTTGAATGCTTCTGATCTACACAGAACTAGAAGATCCATTCCAG 60  
|||||  
Db 795 ATGGAATACTTGAATGCTTCTGATCTACACAGAACTAGAAGATCCATTCCAG 736

QY 61 AACACCTGTAACAGACACCGAGATCTGCTTCTCTGCGGAGCTCGGCGACGAC 120  
|||||  
Db 735 AACACCTGTAACAGACACCGAGATCTGCTTCTCTGCGGAGCTCGGCGACGAC 676

QY 121 TTCTTCTCTCCCGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180  
|||||  
Db 675 TTCTTCTCTCCCGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 616

QY 181 GTCTTGT 240  
|||||  
Db 615 GTCTTGT 556

QY 241 CTCTTGAAGCTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300  
|||||  
Db 555 CTCTTGAAGCTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 496

QY 301 TATGAGATGTGGGCGCAACTACCTTTCTTGTGGGCGCGTGGGCTGCTACTTCAAGAG 360  
|||||  
Db 495 TATGAGATGTGGGCGCAACTACCTTTCTTGTGGGCGCGTGGGCTGCTACTTCAAGAG 436

QY 361 GCCCTCTTTGAGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420  
|||||  
Db 435 GCCCTCTTTGAGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 376

QY 421 CGCTAGTGAGCGATCTGACACCGGCTTCCGCAACTGAGAGACCGGCGCGGCGGC 480  
|||||  
Db 375 CGCTAGTGAGCGATCTGACACCGGCTTCCGCAACTGAGAGACCGGCGCGGCGGC 316

QY 481 CTCAGGATCTCGGCAATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540  
|||||  
Db 315 CTCAGGATCTCGGCAATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 256

QY 541 ATCCATGCGATCAAGTTCACACTTCCCAATGGGCTCCCTGTGCCAGGTTCCGCGACG 600  
|||||  
Db 255 ATCCATGCGATCAAGTTCACACTTCCCAATGGGCTCCCTGTGCCAGGTTCCGCGACG 196

QY 601 TGTAGGTCATCAAGCCCATGTGATCTCAATTTTATCATCTCCAGGTCACCTCTCTCTA 660  
|||||  
Db 195 TGTAGGTCATCAAGCCCATGTGATCTCAATTTTATCATCTCCAGGTCACCTCTCTCTA 136

QY 661 TTCTTACCTCTCCCGATGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 720  
|||||  
Db 135 TTCTTACCTCTCCCGATGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 76

QY 721 AGTATCTAG 729  
|||||  
Db 75 AGTATCTAG 67

RESULT 3  
AC008571/c  
LOCUS AC008571 214267 bp DNA linear PRI 26-SEP-2001  
DEFINITION Homo sapiens chromosome 5 clone cnc-550M4, complete sequence.  
ACCESSION AC008571  
VERSION AC008571.6 GI:15778683  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 214267)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Direct Submission  
TITLE Unpublished  
JOURNAL 2 (bases 1 to 214267)  
DOE Joint Genome Institute.  
REFERENCE Direct Submission  
TITLE Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
JOURNAL 3 (bases 1 to 214267)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Direct Submission  
TITLE Submitted (26-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On Sep 26, 2001 this sequence version replaced gi:15290296.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.8% of Sequence;  
Estimated Total Number of Errors is 0.3.  
FEATURES  
source  
1..214267  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTC-550M4"  
BASE COUNT 62950 a 40311 c 43061 g 67945 t  
ORIGIN  
Query Match 100.0%; Score 729; DB 9; Length 214267;  
Query Local Similarity 100.0%; Pred. No. 1.4e-127;  
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAAAACTTGAATGCTTCTGATCTACACAGAACTAGAACTTCATTCACG 60  
DB 104727 ATGGAAAACTTGAATGCTTCTGATCTACACAGAACTAGAACTTCATTCACG 104668  
QY 61 AAACACCTGAAGCAGCAGGAGTATCGGCTTCTGCGGAGCTCGGCGACCCAC 120  
DB 104667 AAACACCTGAAGCAGCAGGAGTATCGGCTTCTGCGGAGCTCGGCGACCCAC 104608  
QY 121 TTCTTCTCCCGCTGTCTGTGATGTGCAATTTTGTGGGGGATTTGGCAAT 180  
DB 104607 TTCTTCTCCCGCTGTCTGTGATGTGCAATTTTGTGGGGGATTTGGCAAT 104548  
QY 181 GTCTGTGTGCTGTGATTTGACAGACACGAGTATGAAGAGCCCACTACTAC 240  
DB 104547 GTCTGTGTGCTGTGATTTGACAGACACGAGTATGAAGAGCCCACTACTAC 104488  
QY 241 CTCTTACGCTGGCGGTCTGTGACCTCGGCTCTCTCTTGGATGCCCCGAGAGTC 300  
DB 104487 CTCTTACGCTGGCGGTCTGTGACCTCGGCTCTCTCTTGGATGCCCCGAGAGTC 104428  
QY 301 TATGAGATGTGGGCACTACCTTTCTTTGGGGCCGCTGGCTGTACTTCAAGAGC 360  
DB 104427 TATGAGATGTGGGCACTACCTTTCTTTGGGGCCGCTGGCTGTACTTCAAGAGC 104368  
QY 361 GCCCTTTTGAGACGCTGTGCTTGGCTTCATCTCAGATCAACCACTGACGCTGAG 420  
DB 104367 GCCCTTTTGAGACGCTGTGCTTGGCTTCATCTCAGATCAACCACTGACGCTGAG 104308

QY 421 CGCTAGGTGGCCATCGTACACCGCTTCCGGCCCAACTGACAGACACCGCGCGGCC 480  
DB 104307 CGCTAGGTGGCCATCGTACACCGCTTCCGGCCCAACTGACAGACACCGCGCGGCC 104248  
QY 481 CTCAGATCTCGGCAATCGTCTGGGCTTCTCCGCTCTTCTCCGCGCCACACAGC 540  
DB 104247 CTCAGATCTCGGCAATCGTCTGGGCTTCTCCGCTCTTCTCCGCGCCACACAGC 104188  
QY 541 ATCCATGGCATCAAGTTCACATCTTCCCAATGGTCCCTGCTCCAGTTGCGCAC 600  
DB 104187 ATCCATGGCATCAAGTTCACATCTTCCCAATGGTCCCTGCTCCAGTTGCGCAC 104128  
QY 601 TGTAGGTATCAAGCCCATGTGATCTACAAATTTATATATCAGAGTACACTCTTCTA 660  
DB 104127 TGTAGGTATCAAGCCCATGTGATCTACAAATTTATATATCAGAGTACACTCTTCTA 104068  
QY 661 TTCTACCTCTCCCGCATGATCTATCAGTGTCTCTACACTCATGACACAGAGT 720  
DB 104067 TTCTACCTCTCCCGCATGATCTATCAGTGTCTCTACACTCATGACACAGAGT 104008  
QY 721 AGTATCTAG 729  
DB 104007 AGTATCTAG 103999  
RESULT 4  
AF242874  
LOCUS AF242874 1239 bp mRNA linear PRI 13-JUL-2000  
DEFINITION Homo sapiens neuromedin U receptor 2 (NMU2R) mRNA, complete cds.  
ACCESSION AF242874  
VERSION AF242874.1 GI:9082155  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 1239)  
Howard, A.D., Wang, R., Pong, S.S., Mellin, T.N., Strack, A., Guan, X.M.,  
Zeng, Z., Williams, D.L., Jr., Feighner, S.D., Nunes, C.N., Murphy, B.,  
Stall, J.N., Yu, H., Jiang, Q., Clements, M.K., Tan, C.P., McKee, K.K.,  
Hreniuk, D.L., McDonald, T.P., Lynch, K.R., Evans, J.F., Austin, C.P.,  
Caskley, C.T., Van der Ploeg, L.H. and Liu, Q.  
Identification of receptors for neuromedin U and its role in  
feeding  
Nature 406 (6791), 70-74 (2000)  
JOURNAL Nature 406 (6791), 70-74 (2000)  
MEDLINE 20351041  
PUBMED 10894543  
REFERENCE 2 (bases 1 to 1239)  
Liu, Q., McDonald, T.P., Wang, R., Jiang, Q. and Howard, A.D.  
Direct Submission  
Submitted (09-MAR-2000) Pharmacology, Merck Research Labs, West  
Point, PA 19486, USA  
FEATURES  
source  
1..1239  
Location/Qualifiers  
1..1239  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
1..1239  
/gene="NMU2R"  
1..1239  
/gene="NMU2R"  
/note="NMU2R"  
/codon\_start=1  
/product="neuromedin U receptor 2"  
/protein\_id="AAF82755.1"  
/db\_xref="GI:9082156"  
/translation="MEKIDNASMTYQKLEDPFKHINSTEYLAFLCPGRSHFLP  
VSVVPLPVGIVGIVLVCVILIQHQAQKTPNLYLFLSLAVSDLVLLGLPVEYE  
LMRNVPLEGPVGCYFKTALFETVCFASILSTTVSVERVVA1LHPFRKLOSTRRA  
IRLITIVGFSVLFSPNPSIGIKFHFHSPNGSLVPGSATCTVIRPMIYNF1IOVTS  
FLVILPMIVISVILYIMALRLKDKKSLKLEADGNNINIORPSKSVKMLFVILVFAI  
CNAFPHIDRLPFSFVEMESLSLAIFNLVHVVSQVFYLLSSAVNP1IYNLSRRQAA

FQNVISSEPHKQWHSQHDLPOLPPAQRNIFLTBCHFEVLELTDIGPOPCOSSMHNHSLPT  
 ALSSBQMSRTNVSFHNKT"  
 BASE COUNT 267 a 392 c 265 g 315 t  
 ORIGIN  
 Query Match 98.5%; Score 717.8; DB 9; Length 1239;  
 Best Local Similarity 99.7%; Pred. No. 2.2e-125;  
 Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGGAAGAACTTGAATGCTTCTGATCTACACAGAACTAGAAATTCATTCCAG 60
   |||||||
DB 1 ATGGAAGAACTTGAATGCTTCTGATCTACACAGAACTAGAAATTCATTCCAG 60
QY 61 AAACACCTGAACAGACACCGAGAGATATGACCTTCTGCGACCTGCGACGAC 120
   |||||||
DB 61 AAACACCTGAACAGACACCGAGAGATATGACCTTCTGCGACCTGCGACGAC 120
QY 121 TTCTTCTCCCTCCCTGCTGCTGTGATGTCGCAATTTTGTGTGGGGGTCA 180
   |||||||
DB 121 TTCTTCTCCCTCCCTGCTGCTGTGATGTCGCAATTTTGTGTGGGGGTCA 180
QY 181 GTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
   |||||||
DB 181 GTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 CTCTTCAGCGCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
   |||||||
DB 241 CTCTTCAGCGCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 TATGAGATGTGGCGCACTACCTTTCTGCTGCGGCGCTGCTGCTGCTGCT 360
   |||||||
DB 301 TATGAGATGTGGCGCACTACCTTTCTGCTGCGGCGCTGCTGCTGCTGCT 360
QY 361 GCCCTTTGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
   |||||||
DB 361 GCCCTTTGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 CGCTACGCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
   |||||||
DB 421 CGCTACGCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 CTCGAGATCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
   |||||||
DB 481 CTCGAGATCTCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 ATGCATGCGATCAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
   |||||||
DB 541 ATGCATGCGATCAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 TGTACGCTGATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
   |||||||
DB 601 TGTACGCTGATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 TTCTACTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
   |||||||
DB 661 TTCTACTCTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 A 721
DB 721 A 721
  
```

RESULT 5  
 AB041228  
 LOCUS Homo sapiens mRNA for G protein-coupled receptor TGR-1, complete cds.  
 DEFINITION  
 AB041228 1248 bp mRNA linear PRI 21-SEP-2000  
 ACCESSION  
 AB041228  
 VERSION  
 AB041228.1 GI:10257380  
 KEYWORDS  
 G protein-coupled receptor TGR-1.  
 SOURCE  
 Homo sapiens CDNA to mRNA.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (sites)  
 Hosoya,M., Moriya,T., Kawamata,Y., Ohkubo,S., Fujii,R., Matsui,H.,  
 Shintani,Y., Fukusumi,S., Habata,Y., Hinuma,S., Onda,H.,  
 Nishimura,O. and Fujino,M.  
 Identification and functional characterization of a novel subtype  
 of neuromedin U receptor  
 J. Biol. Chem. 275 (38), 29528-29532 (2000)  
 JOURNAL  
 MEDLINE  
 20449029  
 REFERENCE  
 2 (bases 1 to 1248)  
 Shintani,Y., Moriya,T., Ohkubo,S. and Matsui,H.  
 Direct Submision  
 Submitted (03-APR-2000) Yasushi Shintani, Takeda Chemical  
 Industries, Ltd, Discovery Research Laboratories 1; 10 Wadal,  
 Tsukuba, Ibaraki 300-4293, Japan  
 (E-mail:Shintani.Yasushi@takeda.co.jp, Tel:81-298-64-5011,  
 Fax:81-298-64-5000)

FEATURES  
 source  
 Location/Qualifiers  
 1..1248  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1..1248  
 /gene="TGR-1"  
 10..1248  
 /gene="TGR-1"  
 /codon\_start=1  
 /product="G protein-coupled receptor TGR-1"  
 /protein\_id="BAB13721.1"  
 /translation="MEKLNQASWIVYOKLEDPQKLNSTEEYLAFLGPRSRHFLP  
 VSVVVPLEFVGIVNLVLCVILIOHQAMKPTNYLSLAVSDLVLLGMPLEVE  
 MWRNRPLEFVGIVNLVLCVILIOHQAMKPTNYLSLAVSDLVLLGMPLEVE  
 LRILGIVGSEVLSPLPMTSHIGTKEHNPSEIVGATGVICVIRPMVYNFIQVTS  
 FLFLVILGIVGSEVLSPLPMTSHIGTKEHNPSEIVGATGVICVIRPMVYNFIQVTS  
 CWAPHIDRLFFSFVEEMSESLAFAFNIVHYVSGVFYLLSVNITYNLSRRQAA  
 FQNVISSEPHKQWHSQHDLPOLPPAQRNIFLTBCHFEVLELTDIGPOPCOSSMHNHSLPT  
 ALSSBQMSRTNVSFHNKT"  
 BASE COUNT 269 a 393 c 269 g 317 t  
 ORIGIN

Query Match 98.5%; Score 717.8; DB 9; Length 1248;  
 Best Local Similarity 99.7%; Pred. No. 2.2e-125;  
 Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGGAAGAACTTGAATGCTTCTGATCTACACAGAACTAGAAATTCATTCCAG 60
   |||||||
DB 10 ATGGAAGAACTTGAATGCTTCTGATCTACACAGAACTAGAAATTCATTCCAG 69
QY 61 AAACACCTGAACAGACACCGAGAGATATGACCTTCTGCGACCTGCGACGAC 120
   |||||||
DB 70 AAACACCTGAACAGACACCGAGAGATATGACCTTCTGCGACCTGCGACGAC 129
QY 121 TTCTTCTCCCTCCCTGCTGCTGTGATGTCGCAATTTTGTGTGGGGGTCA 180
   |||||||
DB 130 TTCTTCTCCCTCCCTGCTGCTGTGATGTCGCAATTTTGTGTGGGGGTCA 189
QY 181 GTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
   |||||||
DB 190 GTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
QY 241 CTCTTCAGCGCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
   |||||||
DB 250 CTCTTCAGCGCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
QY 301 TATGAGATGTGGCGCACTACCTTTCTGCTGCGGCGCTGCTGCTGCTGCT 360
   |||||||
DB 310 TATGAGATGTGGCGCACTACCTTTCTGCTGCGGCGCTGCTGCTGCTGCT 369
QY 361 GCCCTTTGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
   |||||||
DB 370 GCCCTTTGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
QY 421 CGCTACGCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
  
```

|||||  
Db 430 CGCTAGGTGGCCATCTACACCCGCTCCGGCCAACTGCAGACACCGGGCGCGGGCC 489  
481 CTGAGATCTCGGCATCGCTGCGGGCTTCGCTGCTCTTCTCCCTGCCCAACACAGC 540  
Db 490 CTGAGATCTCGGCATCGCTGCGGGCTTCGCTGCTCTTCTCCCTGCCCAACACAGC 549  
541 ATCCATGGCATCAAGTCTCACTACTTCCCAATGGTCCCTGGTCCAGGTTGGCCACC 600  
Db 550 ATCCATGGCATCAAGTCTCACTACTTCCCAATGGTCCCTGGTCCAGGTTGGCCACC 609  
601 TGTACGGTCATCAAGCCCATGTGATCTCAATTTTCATCATCCAGGTCACCTCTTCTCA 660  
Db 610 TGTACGGTCATCAAGCCCATGTGATCTCAATTTTCATCATCCAGGTCACCTCTTCTCA 669  
661 TTCTACTCTCTCCCATGACTGTGCATCAGTGTCTTACTACTCATGGACATCAAGTGTG 720  
Db 670 TTCTACTCTCTCCCATGACTGTGCATCAGTGTCTTACTACTCATGGACATCAAGTGTG 729  
Oy 721 A 721  
Db 730 A 730  
RESULT 6  
AF272363 1298 bp mRNA linear PRI 23-OCT-2000  
LOCUS Homo sapiens neuromedin U receptor 2 (NMUR2) mRNA, complete cds.  
DEFINITION AF272363  
ACCESSION AF272363.1 GI:10946202  
VERSION  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1298)  
RADATZ,R., WILSON,A.E., ARTYMSHYN,R., BONINI,J.A., BOROWSKY,B.,  
DAIJU,L.W., ZHOU,S., KOURANOVA,E.V., NAGORNY,R., GUEVARA,M.S.,  
DAI,M., LERMAN,G.S., VAYSE,P.J., BRANCHEK,T.A., GERALD,C.,  
FORRAY,C. and ADHAM,N.  
Identification and characterization of two neuromedin U receptors  
differentially expressed in peripheral tissues and the central  
nervous system  
J. Biol. Chem. 275 (42), 32452-32459 (2000)  
PUBMED 10899166  
JOURNAL 2 (bases 1 to 1298)  
EDLINE Bonini,J.A., Radatz,R., Wilson,A. and Borowsky,B.  
20490668 Direct Submission  
AUTHORS Submitted (25-MAY-2000) Target Discovery and Assessment, Synaptic  
JOURNAL Pharmaceutical Corporation, 215 College Road, Paramus, NY 07652,  
USA  
FEATURES  
source Location/Qualifiers  
1..1298  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5q31.1-q31.3"  
1..1298  
/gene="NMUR2"  
27..1274  
/gene="NMUR2"  
/note="G protein-coupled receptor"  
/codon\_start=1  
/product="neuromedin U receptor 2"  
/protein\_id="AA024794.1"  
/db\_xref="GI:10946203"  
/translation="MSGKELDNASMIYQKLEDPFKHNTSEEVLAFLCGPRRSHF  
FLPSVYVPIPVGVIGNVLCIVLIQHQAKPBNLYLFSGLAVSDLIYLLGMPLE  
VYEMRNYPLEGPVGCYEFKTALEFYVCAASLSTTYSVEYVAALHFRKLDKSTR  
RAALRILGVKFSVLFSLPNTSIRKHFIPNSGLVGSATCTVIAKMLYNTLIIQ  
VTSFLYLPVIVISLVLYLALRLKDKSLDADGNAINIQRCKSVKMLFVLYLV  
FAICWAPPHIDRLFPFSVEWSESIAAVFNLVHVVSQVFEYLSAVNPITYMLISRF

GAAPONVLSFHKWHQSDHPOLPPAQRNIFLTCEHFVELTEDIGPQPPCOSSMHNSH  
LPTALISSEMSRNTYOSFHNKT"  
BASE COUNT 278 a 405 c 282 g 333 t  
ORIGIN  
Query Match 98.5%; Score 717.8; DB 9; Length 1298;  
Best local Similarity 99.7%; Pred. No. 2,2e-125;  
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 ATGAAAACCTTGAGATGCTTCTCGATCTACACAGAACTAGAAATCCATTCCAG 60  
Db 36 ATGAAAACCTTGAGATGCTTCTCGATCTACACAGAACTAGAAATCCATTCCAG 95  
Oy 61 AACACCTGAACAGCACCGGAGATATCGGCTTCTCGGAGACTCGGGCCAGCCAC 120  
Db 96 AACACCTGAACAGCACCGGAGATATCGGCTTCTCGGAGACTCGGGCCAGCCAC 155  
Oy 121 TTCTCTCTCCCGCTGTCTGTGTGTATGTGCAATTTTGTGTGGGGGTCAATGGCAAT 180  
Db 156 TTCTCTCTCCCGCTGTCTGTGTGTATGTGCAATTTTGTGTGGGGGTCAATGGCAAT 215  
Oy 181 GTCTGT 240  
Db 216 GTCTGT 275  
Oy 241 CTCTTCAAGCTGGGCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300  
Db 276 CTCTTCAAGCTGGGCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 335  
Oy 301 TATGAGATGTGGCGCAACTACCTTTCTTGTGGGCGCGTGTGTGTGTGTGTGTGTGTGT 360  
Db 336 TATGAGATGTGGCGCAACTACCTTTCTTGTGGGCGCGTGTGTGTGTGTGTGTGTGTGT 395  
Oy 361 GCCCTTTTGACACCGT 420  
Db 396 GCCCTTTTGACACCGT 455  
Oy 421 CGCTAGTGGCCATCTCTACACCGGCTTCCGGCCAACTGAGAGCACCGGCGCGGGCC 480  
Db 456 CGCTAGTGGCCATCTCTACACCGGCTTCCGGCCAACTGAGAGCACCGGCGCGGGCC 515  
Oy 481 CTGAGATCTCTCGGCATCGTCTGCGGCTTCTCCGCTCTTCTCCCTGCCCAACACAGC 540  
Db 516 CTGAGATCTCTCGGCATCGTCTGCGGCTTCTCCGCTCTTCTCCCTGCCCAACACAGC 575  
Oy 541 ATCCATGGCATCAAGTCTCACTACTTCCCAATGGTCCCTGGTCCAGGTTGGCCACC 600  
Db 576 ATCCATGGCATCAAGTCTCACTACTTCCCAATGGTCCCTGGTCCAGGTTGGCCACC 635  
Oy 601 TGTACGGTCATCAAGCCCATGTGATCTCAATTTTCATCATCCAGGTCACCTCTTCTCA 660  
Db 636 TGTACGGTCATCAAGCCCATGTGATCTCAATTTTCATCATCCAGGTCACCTCTTCTCA 695  
Oy 661 TTCTACTCTCTCCCATGACTGTGCATCAGTGTCTTACTACTCATGGACATCAAGTGTG 720  
Db 696 TTCTACTCTCTCCCATGACTGTGCATCAGTGTCTTACTACTCATGGACATCAAGTGTG 755  
Oy 721 A 721  
Db 756 A 756  
RESULT 7  
AX109242 1594 bp DNA linear PAT 30-APR-2001  
LOCUS AX109242  
DEFINITION Sequence 9 from Patent WO0125269.  
ACCESSION AX109242  
VERSION AX109242.1 GI:13924115  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1594)  
 AUTHORS Delaer, J. M., Berger, C., Loeken, C., Nys, G. and Venema, J.  
 TITLE Human g-protein coupled receptor  
 JOURNAL Patent: WO 0125269-A 9 12-APR-2001;  
 Solvay Pharmaceuticals B.V. (NL)  
 FEATURES  
 source Location/Qualifiers  
 .1594  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 55..945  
 /note="IGS4A truncated DNA long version"  
 CDS  
 /protein\_id="CAC37775.1"  
 /db\_xref="GI:13924116"  
 /translation="MSGMKLQNSWYQKLEDPFQKLNSTERYIAFLGPRSHF  
 ELPSVYVPIRVGVIGNVLYCLVILQDAKPTPNLYLSLSDLVLLGMPLE  
 VYEMRNIPFLGPGVGYKIALFEVCAASLSTTVYVERVAILHFRKLQSTR  
 RALRLIGIVGFSVLESLPNTSHGKHPNCSLVGSACTVYIKPMIYNFIIO  
 VSFLEFLPMTVISLYLYLMALRLKDKSLKLEADBNANIORPKRSVNMKLSLRSG  
 VNPMLCSTSSMWCQSSST"  
 BASE COUNT 370 a 473 c 332 g 419 t  
 ORIGIN  
 Query Match 98.5%; Score 717.8; DB 6; Length 1594;  
 Best Local Similarity 99.7%; Pred. No. 2.2e-125;  
 Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 724 TTCTACCTCCTCCCATGACTGTATCATAGTCTCTACTACTCATGCACACTGACACTA 783  
 QY 721 A 721  
 Db 784 A 784  
 RESULT 8  
 AXI09244  
 LOCUS Sequence 11 from Patent WO0125269.  
 DEFINITION AXI09244  
 ACCESSION AXI09244  
 VERSION AXI09244.1 GI:13924117  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 1594)  
 AUTHORS Delaer, J. M., Berger, C., Loeken, C., Nys, G. and Venema, J.  
 TITLE Human g-protein coupled receptor  
 JOURNAL Patent: WO 0125269-A 11 12-APR-2001;  
 Solvay Pharmaceuticals B.V. (NL)  
 FEATURES  
 source Location/Qualifiers  
 .1594  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 64..945  
 /note="IGS4A truncated DNA short version"  
 CDS  
 /protein\_id="CAC37776.1"  
 /db\_xref="GI:13924118"  
 /translation="MEKIQNSWYQKLEDPFQKLNSTERYIAFLGPRSHFELP  
 VSVYVPIRVGVIGNVLYCLVILQDAKPTPNLYLSLSDLVLLGMPLE  
 VYEMRNIPFLGPGVGYKIALFEVCAASLSTTVYVERVAILHFRKLQSTR  
 RALRLIGIVGFSVLESLPNTSHGKHPNCSLVGSACTVYIKPMIYNFIIO  
 VSFLEFLPMTVISLYLYLMALRLKDKSLKLEADBNANIORPKRSVNMKLSLRSG  
 VNPMLCSTSSMWCQSSST"  
 BASE COUNT 370 a 473 c 332 g 419 t  
 ORIGIN  
 Query Match 98.5%; Score 717.8; DB 6; Length 1594;  
 Best Local Similarity 99.7%; Pred. No. 2.2e-125;  
 Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



QY 421 CGTACGTGGCCATCTACACCCGTTCCGGCCAAACTGCAGACACCGCGCGGGCC 480  
|||||  
Db 484 CGTACGTGGCCATCTACACCCGTTCCGGCCAAACTGCAGACACCGCGCGGGCC 543  
QY 481 CTGAGGATCTCGGCATCGTCTGGGGCTTCTCCGTCTCTTCTCCCTGCCCAACACGAGC 540  
|||||  
Db 544 CTGAGGATCTCGGCATCGTCTGGGGCTTCTCCGTCTCTTCTCCCTGCCCAACACGAGC 603  
QY 541 ATCCATGGCATCAAGTTCCTCACTATTCCTCCCAATGGGCTCCCTGGGCCAGGTTGGGCCACC 600  
|||||  
Db 604 ATCCATGGCATCAAGTTCCTCACTATTCCTCCCAATGGGCTCCCTGGGCCAGGTTGGGCCACC 663  
QY 601 TGTACGGTCAATCAAGCCCATGTGGATCTACAAATTCATCTCAGAGTCAACCTCTCTCTTA 660  
|||||  
664 TGTACGGTCAATCAAGCCCATGTGGATCTACAAATTCATCTCAGAGTCAACCTCTCTCTTA 723  
661 TTCTACTCTCTCCCATGATGTCTCATGAGTGTCTCTACTACTCTCATGACATCAGAGTGTG 720  
|||||  
Db 724 TTCTACTCTCTCCCATGATGTCTCATGAGTGTCTCTACTACTCTCATGACATCAGAGTGTG 783  
QY 721 A 721  
|  
Db 784 A 784

RESULT 9  
AX109234 1658 bp DNA linear PAT 30-Apr-2001  
LOCUS Sequence 1 from Patent WO0125269.  
DEFINITION AX109234  
ACCESSION AX109234  
VERSION AX109234.1 GI:13924107  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1658)  
AUTHORS Deleersnijder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.  
TITLE Human g-protein coupled receptor  
JOURNAL Patent: WO 0125269-A 12-Apr-2001;  
Solvay Pharmaceuticals B.V. (NL)  
FEATURES  
source  
location/Qualifiers  
1..1658  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
55..1302  
/note="IGS4A long version"  
CDS  
/codon\_start=1  
/protein\_id="CAC37771.1"  
/db\_xref="GI:13924108"  
/translation="MSGMEKLONASWYIOOKLEDPFOKHLNSTEYLAFLCGPRRSHF  
PLPVSVYPIFVGVIGNVLCVILIOHAKPTPNYLFSLAVSDLVLLGMPLE  
VEEMRNYPFLFGPVCGYFKTALFETVCFASILSTIVSVERVAIILHPRAKLOSTR  
RKALRLGIWGEISVLSLPNTSINGIKRHPNGSLVGSATCTYIKPMWYNTLITQ  
VSWFLEYLLPMYIVLSVLYLMLRLKDKLSLEDEGNANVIOGRKRSVKNMFLVLLV  
FAISGAPFHIDRLFEFSVEEMSESLAENVLHVYGFVFEYSSAVNPITLYNLSRF  
OAFONVYSFHKQSHOHDPOLPQNRNIFLECHFEVLTEDIGPFCQSSMHNHS  
LPTALSSBQMSRNTYOSRHPNKT"

BASE COUNT 377 a 490 c 346 g 445 t  
ORIGIN  
Query Match 98.5%; Score 717.8; DB 6; Length 1658;  
Best local Similarity 99.7%; Pred. No. 2, 2e-125;  
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAAAACTTGAGATGCTTCTGGATCTACACGAGAACTAGAAATCCATTCCAG 60  
|||||  
Db 64 ATGGAAAACTTGAGATGCTTCTGGATCTACACGAGAACTAGAAATCCATTCCAG 123  
QY 61 AAACACCTGAACGACCGGAGATATCTGGCTTCTCTCGGAGCTCGGCGACGCAC 120  
|||||  
Db 124 AAACACCTGAACGACCGGAGATATCTGGCTTCTCTCGGAGCTCGGCGACGCAC 183

QY 121 TTCTCTCCCGCGTCTCTGTGTGTATGTGCCAATTTTGTGTGGGGCTATTTGGCAAT 180  
|||||  
Db 184 TTCTCTCCCGCGTCTCTGTGTGTATGTGCCAATTTTGTGTGGGGCTATTTGGCAAT 243  
QY 181 GTCTGTGTGTCTGTGTATTTGTGCAGACACCAAGGCTATGAAGAGCCCACTACTAC 240  
|||||  
Db 244 GTCTGTGTGTCTGTGTATTTGTGCAGACACCAAGGCTATGAAGAGCCCACTACTAC 303  
QY 241 CTCTTCAGCCGGGGGGCTGTGCACCTCTGTGCTCTCTCTGTGAATGGCCCTGGAGTTC 300  
|||||  
Db 304 CTCTTCAGCCGGGGGGCTGTGCACCTCTGTGCTCTCTCTGTGAATGGCCCTGGAGTTC 363  
QY 301 TATGAGATGTGGCGCAACTACCTCTTCTGTGGCGCCCTGGGGCTCTACTTCAAGAGC 360  
|||||  
Db 364 TATGAGATGTGGCGCAACTACCTCTTCTGTGGCGCCCTGGGGCTCTACTTCAAGAGC 423  
QY 361 GCCCTCTTTGAGACCGTGTCTTGCCTGCATCTCAGATCAGACACCGTACGGTGGAG 420  
|||||  
Db 424 GCCCTCTTTGAGACCGTGTCTTGCCTGCATCTCAGATCAGACACCGTACGGTGGAG 483  
QY 421 CGTACGTGGCCATCTACACCCGTTCCGGCCAAACTGCAGACACCGCGCGGGCC 480  
|||||  
Db 484 CGTACGTGGCCATCTACACCCGTTCCGGCCAAACTGCAGACACCGCGCGGGCC 543  
QY 481 CTGAGATCTCGGCATCGTCTGGGGCTTCCGCTCTTCTCCCTGCCCAACACGAGC 540  
|||||  
Db 544 CTGAGATCTCGGCATCGTCTGGGGCTTCCGCTCTTCTCCCTGCCCAACACGAGC 603  
QY 541 ATCCATGGCATCAAGTTCCTCACTATTCCTCCCAATGGGCTCCCTGGGCCAGGTTGGGCCACC 600  
|||||  
Db 604 ATCCATGGCATCAAGTTCCTCACTATTCCTCCCAATGGGCTCCCTGGGCCAGGTTGGGCCACC 663  
QY 601 TGTACGGTCAATCAAGCCCATGTGGATCTACAAATTCATCTCAGAGTCAACCTCTCTCTTA 660  
|||||  
Db 664 TGTACGGTCAATCAAGCCCATGTGGATCTACAAATTCATCTCAGAGTCAACCTCTCTCTTA 723  
QY 661 TTCTACTCTCTCCCATGATGTCTCATGAGTGTCTCTACTACTCTCATGACATCAGAGTGTG 720  
|||||  
Db 724 TTCTACTCTCTCCCATGATGTCTCATGAGTGTCTCTACTACTCTCATGACATCAGAGTGTG 783  
QY 721 A 721  
|  
Db 784 A 784

RESULT 10  
AX109236 1658 bp DNA linear PAT 30-Apr-2001  
LOCUS Sequence 3 from Patent WO0125269.  
DEFINITION AX109236  
ACCESSION AX109236  
VERSION AX109236.1 GI:13924109  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1658)  
AUTHORS Deleersnijder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.  
TITLE Human g-protein coupled receptor  
JOURNAL Patent: WO 0125269-A 3 12-Apr-2001;  
Solvay Pharmaceuticals B.V. (NL)  
FEATURES  
source  
location/Qualifiers  
1..1658  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
64..1302  
/note="IGS4A short version"  
CDS  
/codon\_start=1  
/protein\_id="CAC37772.1"  
/db\_xref="GI:13924110"  
/translation="MEKLNASWYIOOKLEDPFOKHLNSTEYLAFLCGPRRSHFLP  
VSVYVPIFVGVIGNVLCVILIOHAKPTPNYLFSLAVSDLVLLGMPLEVEE  
MERNYPFLFGPVCGYFKTALFETVCFASILSTIVSVERVAIILHPRAKLOSTRRA

LRILGIWGFVSLEPNTSINGIKRHPNGLVDSATCTVIKPMIYNFIQVTS  
FLFYLLPMVIVISLVYLMALRLKDKSLDEADENANIORCRSVKMLFVLVFAI  
CMAPFHIDRLFEFVEEMTSLAAVNLVHVVSGLYFLSSAVNP1IYNLSRPOAA  
FONVSSFHOMHSDHDPOLPPAORNIPLTECHFVELTEDIGFPOQSSVHNSHLP  
ALSSEOMSRNINOSFHNKT"

BASE COUNT 377 a 490 c 346 g 445 t

Query Match 98.5%; Score 717.8; DB 6; Length 1658;  
Best Local Similarity 99.7%; Pred. No. 2.2e-125;  
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAATACTTGAATGCTTCTGATCTACACGAGAAACTAGAATCCATTCCAG 60  
|||||  
Db ATGGAATACTTGAATGCTTCTGATCTACACGAGAAACTAGAATCCATTCCAG 123  
QY 61 AAACACTGAAACGACGAGAGATATCTGGCTTCTCTGCGAAGCTCGGCGAGCCAC 120  
|||||  
Db AAACACTGAAACGACGAGAGATATCTGGCTTCTCTGCGAAGCTCGGCGAGCCAC 183  
QY 121 TTCTTCTCCCGCTGCTGTGTATGTATGTGCAATTTTGGTGGGGGTCATTTGGCAAT 180  
|||||  
Db TTCTTCTCCCGCTGCTGTGTATGTATGTGCAATTTTGGTGGGGGTCATTTGGCAAT 243  
QY 181 GTCTGTGTGTGCTGTGTATGTGCAACACGAGCTATGAAGAGCCGCAACTACTAC 240  
|||||  
Db GTCTGTGTGTGCTGTGTATGTGCAACACGAGCTATGAAGAGCCGCAACTACTAC 303  
QY 241 CTCTTACGCTGCGGGGCTGTGACCTCTGCTGCTCTTGGAAATGCCCTGAGGTC 300  
|||||  
Db CTCTTACGCTGCGGGGCTGTGACCTCTGCTGCTCTTGGAAATGCCCTGAGGTC 363  
QY 301 TATGAGATGTGGGCACTACCTCTTGTGCGGCGCGGCTGCTACTTCAAGACG 360  
|||||  
Db TATGAGATGTGGGCACTACCTCTTGTGCGGCGCGGCTGCTACTTCAAGACG 423  
QY 361 GCCCTCTTGAAGACGCTGTGCTTGGCTCATCTTCAAGATCAACACGCTGAGG 420  
|||||  
Db GCCCTCTTGAAGACGCTGTGCTTGGCTCATCTTCAAGATCAACACGCTGAGG 483  
QY 421 CGCTAGCTGGCAATCCATCAACCGCTTCCGCGCAAAATGAGAGACACCGCGCGGCC 480  
|||||  
Db CGCTAGCTGGCAATCCATCAACCGCTTCCGCGCAAAATGAGAGACACCGCGCGGCC 543  
QY 481 CTCAGATCTCTGCGATCTGTGCGGCTTCTCGTGTCTTCTCCGCGCAACACGAG 540  
|||||  
Db CTCAGATCTCTGCGATCTGTGCGGCTTCTCGTGTCTTCTCCGCGCAACACGAG 603  
QY 541 ATCCATGAGCATCAAGTTCCACTACTTCCCAATGGGTCCTGCTGCCAGGTTCCGACCC 600  
|||||  
Db ATCCATGAGCATCAAGTTCCACTACTTCCCAATGGGTCCTGCTGCCAGGTTCCGACCC 663  
QY 601 TGTACGGTCAATCAAGCCATGTGATCTCAATTTGATCATCATGAGTCACTCTTCTTA 660  
|||||  
Db TGTACGGTCAATCAAGCCATGTGATCTCAATTTGATCATCATGAGTCACTCTTCTTA 723  
QY 661 TTCTACCTCTCCCATGACTGTATCATGTCCTTACTACTCAATGAGCATCAAGT 720  
|||||  
Db TTCTACCTCTCCCATGACTGTATCATGTCCTTACTACTCAATGAGCATCAAGT 783  
QY 721 A 721  
Db 784 A 784

RESULT 11  
AX109238 1658 bp DNA linear PAT 30-APR-2001  
LOCUS AX109238  
DEFINITION Sequence 5 from Patent WO0125269.  
ACCESSION AX109238  
VERSION AX109238.1 GI:13924111  
KEYWORDS  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1658)  
AUTHORS Deleersnijder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.  
TITLE Human g-protein coupled receptor  
JOURNAL Patent: WO 0125269-A 5 12-APR-2001;  
Solvey Pharmaceuticals B.V. (NL)

FEATURES  
Source location/Qualifiers  
1..1658  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
55..1302  
/note="TGS4B long version"  
/protein\_start=1  
/protein\_id="CAC37773.1"  
/db\_xref="GI:13924112"  
/translation="MSGMKLONASWVIOOKLEDDPFOKLHNSTEYIAFLGPRRSHF  
FLPVSVYVPIFVGVIGNVLCIVTIOHAKMTPNTNYFLSLAVSDLLVILGMPLE  
VYEMRNYPFLGPGVGCYFKALFETVCFASLSTITVSVEKRYALHFRKLQSTR  
RRALRLGIWGFVSLEPNTSINGIKRHPNGLVDSATCTVIKPMIYNFIQVTS  
FLFYLLPMVIVISLVYLMALRLKDKSLDEADENANIORCRSVKMLFVLVFAI  
CMAPFHIDRLFEFVEEMTSLAAVNLVHVVSGLYFLSSAVNP1IYNLSRPF  
FONVSSFHOMHSDHDPOLPPAORNIPLTECHFVELTEDIGFPOQSSVHNSH  
LPTALSSEOMSRNINOSFHNKT"

BASE COUNT 378 a 489 c 346 g 445 t

Query Match 98.5%; Score 717.8; DB 6; Length 1658;  
Best Local Similarity 99.7%; Pred. No. 2.2e-125;  
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAATACTTGAATGCTTCTGATCTACACGAGAAACTAGAATCCATTCCAG 60  
|||||  
Db ATGGAATACTTGAATGCTTCTGATCTACACGAGAAACTAGAATCCATTCCAG 123  
QY 61 AAACACTGAAACGACGAGAGATATCTGGCTTCTCTGCGAAGCTCGGCGAGCCAC 120  
|||||  
Db AAACACTGAAACGACGAGAGATATCTGGCTTCTCTGCGAAGCTCGGCGAGCCAC 183  
QY 121 TTCTTCTCCCGCTGCTGTGTATGTGCAATTTTGGTGGGGGTCATTTGGCAAT 180  
|||||  
Db TTCTTCTCCCGCTGCTGTGTATGTGCAATTTTGGTGGGGGTCATTTGGCAAT 243  
QY 181 GTCTGTGTGTGCTGTGTATGTGCAACACGAGCTATGAAGAGCCGCAACTACTAC 240  
|||||  
Db GTCTGTGTGTGCTGTGTATGTGCAACACGAGCTATGAAGAGCCGCAACTACTAC 303  
QY 241 CTCTTACGCTGCGGGGCTGTGACCTCTGCTGCTCTTGGAAATGCCCTGAGGTC 300  
|||||  
Db CTCTTACGCTGCGGGGCTGTGACCTCTGCTGCTCTTGGAAATGCCCTGAGGTC 363  
QY 301 TATGAGATGTGGGCACTACCTCTTGTGCGGCGCGGCTGCTACTTCAAGACG 360  
|||||  
Db TATGAGATGTGGGCACTACCTCTTGTGCGGCGCGGCTGCTACTTCAAGACG 423  
QY 361 GCCCTCTTGAAGACGCTGTGCTTGGCTCATCTTCAAGATCAACACGCTGAGG 420  
|||||  
Db GCCCTCTTGAAGACGCTGTGCTTGGCTCATCTTCAAGATCAACACGCTGAGG 483  
QY 421 CGCTAGCTGGCAATCCATCAACCGCTTCCGCGCAAAATGAGAGACACCGCGCGGCC 480  
|||||  
Db CGCTAGCTGGCAATCCATCAACCGCTTCCGCGCAAAATGAGAGACACCGCGCGGCC 543  
QY 481 CTCAGATCTCTGCGATCTGTGCGGCTTCTCGTGTCTTCTCCGCGCAACACGAG 540  
|||||  
Db CTCAGATCTCTGCGATCTGTGCGGCTTCTCGTGTCTTCTCCGCGCAACACGAG 603  
QY 541 ATCCATGAGCATCAAGTTCCACTACTTCCCAATGGGTCCTGCTGCCAGGTTCCGACCC 600  
|||||  
Db ATCCATGAGCATCAAGTTCCACTACTTCCCAATGGGTCCTGCTGCCAGGTTCCGACCC 663  
QY 601 TGTACGGTCAATCAAGCCATGTGATCTCAATTTGATCATCATGAGTCACTCTTCTTA 660

|||||  
Db 664 TGTACGTCATCAAGCCGATGTGATCATCAATTCATCATCCAGGTCACCTCTCTTA 723  
QY 661 TTCTACCTCTCCCATGACTGTGATCATGTCCTACTACCTCATNGACATCAGTG 720  
Db 724 TTCTACCTCTCCCATGACTGTGATCATGTCCTCTACTACTCATGAGCATCAGACTA 783  
QY 721 A 721  
Db 784 A 784  
RESULT 12  
AXI09240 1658 bp DNA linear PAT 30-APR-2001  
SEQUENCE 7 from Patent WO0125269.  
AXI09240  
VERSION AXI09240.1 GI:13924113  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1658)  
AUTHORS Deleersnijder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.  
TITLE Human g-protein coupled receptor  
JOURNAL Patent: WO 0125269-A 7 12-APR-2001;  
Solvay Pharmaceuticals B.V. (NL)  
FEATURES  
source Location/Qualifiers  
1..1658  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
64..1302  
/note="IGS4B short version"  
/codon\_start=1  
/protein\_id="CAC37774.1"  
/db\_xref="GI:13924114"  
/translation="MEKIQNASMIYQOKLEDPFOKHLNSTEELAFICGPRSHFPLP  
VSVVYPIPVGVGICVILQIQAOKPTPNYYLSLAVSDILVLLGLMPLVEE  
MRNRYBELGCPGYCFKTALEFVCPASLSTTVSVERVAIIHFRAKKLOSTRRA  
LHILGIWGFSVLSPNTSIIHGKIFHPNSGLVPGSATCTVIRKMYLNFIIQVTS  
FLFYLLPMIVISLVLYLMAIRLKKDSLADSGNANIORCKRSVNMKMFVILVFAI  
CWAPFIDILFFSFVEEMTESLAIVENLVHVGSLVFLYSANPTIYMLLSRFQAA  
FONVIFSFKQWMSOHPDLPAPQRIEFLTECHFVELLDIGPFLCQSSVHNSHLP  
ALSSDMSRTNNSFHEKNT"  
BASE COUNT 378 a 489 c 346 g 445 t  
Query Match 98.5%; Score 717.8; DB 6; Length 1658;  
Best local Similarity 99.7%; Pred. No. 2..2e-125;  
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

|||||  
Db 364 TATGAGATGTGGCGCAACTACCTTCTGTGTGGGCCGCTGGGCTGCTACTTCAAGACG 423  
QY 361 GCCCTTTTGAGACCGTGTGCTTGGCTCATCTCAGATFACACACCGGACGTTGGAG 420  
Db 424 GCCCTTTTGAGACCGTGTGCTTGGCTCATCTCAGATFACACACCGGACGTTGGAG 483  
QY 421 CGCTACGTGGCCATCTACACCGCTTCCGGCCAAACTGACAGACACCGGCGGCGC 480  
Db 484 CGCTACGTGGCCATCTACACCGCTTCCGGCCAAACTGACAGACACCGGCGGCGC 543  
QY 481 CTCGAGATCTGGGATCTGTGGGGCTTCCGCTCTTCTCCCTGCCCCAACACACG 540  
Db 544 CTCGAGATCTCGGATCTGTGGGGCTTCCGCTCTTCTCCCTGCCCCAACACACG 603  
QY 541 ATCCATGGCATCAAGTTCACCTTCCCAATGGGTCCTGGTCCAGGTTGGGCGACC 600  
Db 604 ATCCATGGCATCAAGTTCACCTTCCCAATGGGTCCTGGTCCAGGTTGGGCGACC 663  
QY 601 TGTACGTCATCAAGCCGATGTGATCATCAATTCATCAATCCAGGTCACCTCTCTTA 660  
Db 664 TGTACGTCATCAAGCCGATGTGATCATCAATTCATCAATCCAGGTCACCTCTCTTA 723  
QY 661 TTCTACCTCTCCCATGACTGTGATCATGTCCTCTACTACCTCATGCGACATCAGTG 720  
Db 724 TTCTACCTCTCCCATGACTGTGATCATGTCCTCTACTACTCATGCGACATCAGACTA 783  
QY 721 A 721  
Db 784 A 784  
RESULT 13  
AF292402 1248 bp mRNA linear PRI 11-DEC-2000  
LOCUS AF292402  
DEFINITION Homo sapiens neuromedin u receptor-type 2 mRNA, complete cds.  
ACCESSION AF292402  
VERSION AF292402.1 GI:9944989  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1248)  
AUTHORS Shan, L., Qiao, X., Crona, J.H., Behan, J., Wang, S., Laz, T., Bayne, M.,  
Gustafson, E.L., Monsma, F.J., Jr. and Hedrick, J.A.  
TITLE Identification of a novel neuromedin u receptor subtype expressed  
JOURNAL in the central nervous system  
J. Biol. Chem. 275 (50), 39482-39486 (2000)  
MEDLINE 20564321  
PUBMED 11010960  
REFERENCE 2 (bases 1 to 1248)  
AUTHORS Pang, L., Wang, S., Laz, T. and Hedrick, J.A.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2000) Human Genome Research, Schering-Plough  
Research Institute, 2015 Gallloping Hill Road, Kenilworth, NJ 07033,  
USA  
FEATURES  
source Location/Qualifiers  
1..1248  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
1..1248  
/note="G protein coupled receptor; Nmu-R2"  
/codon\_start=1  
/product="neuromedin u receptor-type 2"  
/protein\_id="AA03064.1"  
/db\_xref="GI:9944989"  
/translation="MSGMEKIQNASMIYQOKLEDPFOKHLNSTEELAFICGPRSHF  
FLPVSIVYPIPVGVGICVILQIQAOKPTPNYYLSLAVSDILVLLGLMPLVEE  
VTEBMKNYPLFLPGVGCYFKTALEFVCPASLSTTVSVERVAIIHFRAKKLOSTRRA  
LHILGIWGFSVLSPNTSIIHGKIFHPNSGLVPGSATCTVIRKMYLNFIIQVTS  
FLFYLLPMIVISLVLYLMAIRLKKDSLADSGNANIORCKRSVNMKMFVILVFAI  
CWAPFIDILFFSFVEEMTESLAIVENLVHVGSLVFLYSANPTIYMLLSRFQAA  
FONVIFSFKQWMSOHPDLPAPQRIEFLTECHFVELLDIGPFLCQSSVHNSHLP

FAICWAPFHIDLPFSPVEMSESLAAVENLVHVVGFVYLSSAVNPITYNLISRRF  
 OAAPONVISSEPHKOWHSDHDPOLPPAORNIIFLTECHFVLETFIDIPQPPCCSSMNSH  
 LPTALSSEDMSTRNOSFHFNKT"

BASE COUNT 268 a 393 c 270 g 317 t  
 ORIGIN

Query Match 98.2% Score 716.2; DB 9: Length 1248;  
 Best Local Similarity 99.6%; Pred. No. 4.5e-125;  
 Matches 718; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAAGAACTTCAGAACTTCCTCGATCTACACAGAACTAGAGATCCATTCCAG 60  
 |||||||  
 DB 10 ATGGAAGAACTTCAGAACTTCCTCGATCTACACAGAACTAGAGATCCATTCCAG 69  
 |||||||  
 QY 61 AAACACCTGAACAGACCGAGAGATATCTGGCTTCTCTGCGAGCTGGCGACCCAC 120  
 |||||||  
 DB 70 AAACACCTGAACAGACCGAGAGATATCTGGCTTCTCTGCGAGCTGGCGACCCAC 129  
 |||||||  
 QY 121 TTCTTCCTCCCGGTGTCTGTGGGTATGTGCCAATTTTGTGTGGGGGTCAATGGCAAT 180  
 |||||||  
 DB 130 TTCTTCCTCCCGGTGTCTGTGGGTATGTGCCAATTTTGTGTGGGGGTCAATGGCAAT 189  
 |||||||  
 QY 181 GTCCCTGTGTGCTGTGATTTCTGACAGACAGGCTATGAGAGCCCACTACTATAC 240  
 |||||||  
 DB 190 GTCCCTGTGTGCTGTGATTTCTGACAGACAGGCTATGAGAGCCCACTACTATAC 249  
 |||||||  
 QY 241 CTCTTACGCTGGCGGTCTCTGACCTCCCTGGTCTCTCTTGGAAATGCCCTTGGAGGTC 300  
 |||||||  
 DB 250 CTCTTACGCTGGCGGTCTCTGACCTCCCTGGTCTCTCTTGGAAATGCCCTTGGAGGTC 309  
 |||||||  
 QY 301 TATGAGATGTGGGGGCAACTACCTTTCTGTTGGGGCCGTGGGCTATCTTCAAGAGC 360  
 |||||||  
 DB 310 TATGAGATGTGGGGGCAACTACCTTTCTGTTGGGGCCGTGGGCTATCTTCAAGAGC 369  
 |||||||  
 QY 361 GCCCTCTTGTGAGACCGTGTGCTTGGCTCCATCCCTCAGATCAACACCGTACGCTGAG 420  
 |||||||  
 DB 370 GCCCTCTTGTGAGACCGTGTGCTTGGCTCCATCCCTCAGATCAACACCGTACGCTGAG 429  
 |||||||  
 QY 421 CGCTACGTGGGCAATCTTACACCCGTTCCGGGCAAACTGACAGACCCGGCCGGCC 480  
 |||||||  
 DB 430 CGCTACGTGGGCAATCTTACACCCGTTCCGGGCAAACTGACAGACCCGGCCGGCC 489  
 |||||||  
 QY 481 CTCAGATCTCGGCGCATCTGCTGGGGCTTCCGCTGCTCTCCCTGCCCAACACAGC 540  
 |||||||  
 DB 490 CTCAGATCTCGGCGCATCTGCTGGGGCTTCCGCTGCTCTCTCCCTGCCCAACACAGC 549  
 |||||||  
 QY 541 ATCCATGGCATCAAGTTCACATCTTCCCAATGAGTCCCTGGTCCAGGTTCCGGCCACC 600  
 |||||||  
 DB 550 ATCCATGGCATCAAGTTCACATCTTCCCAATGAGTCCCTGGTCCAGGTTCCGGCCACC 609  
 |||||||  
 QY 601 TGTACGGTATCAAGCCCATGTGATCTACAAATTCATCATCCAGGTCACCTCTTCCTA 660  
 |||||||  
 DB 610 TGTACGGTATCAAGCCCATGTGATCTACAAATTCATCATCCAGGTCACCTCTTCCTA 669  
 |||||||  
 QY 661 TTCTTACCTCTCCCGCATGACTGTGATCTGATCTCTTACACTCAATGGAGCTCAGAGT 720  
 |||||||  
 DB 670 TTCTTACCTCTCCCGCATGACTGTGATCTGATCTCTTACACTCAATGGAGCTCAGAGT 729  
 |||||||  
 QY 721 A 721  
 ||  
 DB 730 A 730  
 ||

RESULT 14  
 LOCUS BC016938 2069 bp mRNA linear PRI 09-NOV-2001  
 DEFINITION Homo sapiens, neuromedin u receptor 2, clone MGC:21396  
 ACCESSION BC016938  
 VERSION BC016938.1 GI:16877376  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens

# REFERENCE

AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 CONTACT: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 CONTACT: (Dickson, Mark) [mdedpaxil.stanford.edu](mailto:mdedpaxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

# REMARK

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 20 Row: n Column: 3  
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.  
 Location/Qualifiers  
 1..2069  
 /organism="Homo sapiens"  
 /db\_xref="locusID:56923"  
 /db\_xref="taxon:9606"  
 /clone="MGC:21396 IMAGE:3852151"  
 /tissue\_type="Colon, adenocarcinoma"  
 /clone\_lib="NIH\_MGC\_65"  
 /lab\_host="DH10B"  
 /note="vector: pCMV-Sport6"  
 162..1409  
 /codon\_start=1  
 /product="neuromedin u receptor 2"  
 /protein\_id="AAH16938.1"  
 /db\_xref="GI:16877377"  
 /translation="MSGMKELONASMIYQKLEDPFOKLNSTEEYIAFLGPRSHFPLPSVYPIPIPVGVIGIVLYLQHQAKTFPNLYLBSLAYSDLYLLGMPLEVEMKNIPFLEPGVCFKTALEFVCHASLITSTVSERYVALHLPFRKLQSTIRRALRILGLVWGSVLSLPTNSHIGIKRHPNPSLVPGSATFVIRKPMIYFNLIQVTSFLFYLIPMTIVSLYLMALRLKDKSLSDAGSNANIQRCKRSVKMLFVLVFAICWAPFHIDLPFSPVEMSESLAAVENLVHVVGFVYLSSAVNPITYNLISRRFOAARONVISSEPHKOWHSDHDPOLPPAORNIIFLTECHFVLETFIDIPQPPCCSSMNSHLPTALSSEDMSTRNOSFHFNKT"

# FEATURES

source

# CDS

BASE COUNT 481 a 579 c 411 g 598 t  
 ORIGIN  
 Query Match 98.2% Score 716.2; DB 9: Length 2069;  
 Best Local Similarity 99.6%; Pred. No. 4.4e-125;  
 Matches 718; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGAAGAACTTCAGAACTTCCTCGATCTACACAGAACTAGAGATCCATTCCAG 60  
 |||||||  
 DB 171 ATGGAAGAACTTCAGAACTTCCTCGATCTACACAGAACTAGAGATCCATTCCAG 220  
 |||||||  
 QY 61 AAACACCTGAACAGACCGAGAGATATCTGGCTTCTCTGCGAGCTGGCGACCCAC 120  
 |||||||  
 DB 231 AAACACCTGAACAGACCGAGAGATATCTGGCTTCTCTGCGAGCTGGCGACCCAC 290  
 |||||||  
 QY 121 TTCTTCCTCCCGGTGTCTGTGGGTATGTGCCAATTTTGTGTGGGGGTCAATGGCAAT 180  
 |||||||  
 DB 291 TTCTTCCTCCCGGTGTCTGTGGGTATGTGCCAATTTTGTGTGGGGGTCAATGGCAAT 350  
 |||||||  
 QY 181 GTCCCTGTGTGCTGTGATTTCTGACAGACAGGCTATGAGAGCCCACTACTATAC 240  
 |||||||  
 DB 351 GTCCCTGTGTGCTGTGATTTCTGACAGACAGGCTATGAGAGCCCACTACTATAC 410  
 |||||||





OY 721 AGTATCTA 728  
|||||  
Db 128570 AGTATCGA 128563

Search completed: January 17, 2003, 04:30:06  
Job time : 3043 secs

THIS PAGE BLANK (4-3870)